

**ORAL STOP CONTRASTS IN OMAHA:  
AN ACOUSTIC ANALYSIS**

by

Elizabeth Ann Neilson

A thesis submitted to the faculty of  
The University of Utah  
in partial fulfillment of the requirements for the degree of

Master of Arts

Department of Linguistics

The University of Utah

August 2011

Copyright © Elizabeth Ann Neilson 2011

All Rights Reserved

# The University of Utah Graduate School

## STATEMENT OF THESIS APPROVAL

The thesis of **Elizabeth Ann Neilson**

has been approved by the following supervisory committee members:

**Lyle Campbell**

, Chair

**March  
13, 2011**

Date Approved

**Ardis Eschenberg**

, Member

**March  
13, 2011**

Date Approved

**Rachel Hayes-Harb**

, Member

**March  
13, 2011**

Date Approved

and by

**Edward Rubin**

, Chair of

the Department of

**Linguistics**

and by Charles A. Wight, Dean of The Graduate School.

## **ABSTRACT**

The goal of the present study is to measure voice onset time (VOT) in the oral stop series in Omaha, a North American Indian language of the Siouan family. The question addressed is: Does acoustic analysis using modern instrumentation support that the phonetic stop categories in Omaha are voiced, voiceless unaspirated, and voiceless aspirated at bilabial, dental, and velar places of articulation as documented in Dorsey's Dhegiha language transcriptions (c. 1890) and attested by other linguists? Then, provided that for each place of articulation, three distinct groups of VOT values form, do these values correlate with VOTs for other languages which contain the three categories listed above, or to other categories of stops? Finally, how can this information be used to improve existing Omaha language revitalization programs and Omaha language lessons? Results showed that even with a variety of cross analyses, there is evidence that the purported three way contrast does consist of voiced, voiceless unaspirated, and voiceless aspirated categories in contemporary Omaha. This can be applied to the Omaha curriculum as a pronunciation guide for students and non-native speaking teachers of Omaha.

## TABLE OF CONTENTS

ABSTRACT.....	iii
ACKNOWLEDGEMENTS.....	v
Chapters	
I. BACKGROUND.....	1
Documenting the world's speech sounds.....	1
Second language acquisition, speech perception, and the field linguist.....	2
VOT across languages.....	4
Omaha language: documentation.....	5
Omaha language: revitalization.....	7
Overview.....	8
II. HYPOTHESIS.....	9
III. METHODS.....	10
Stimuli.....	10
Subjects.....	12
Data collection.....	13
IV. RESULTS/ANALYSIS.....	14
V. DISCUSSION.....	111
Future research.....	120
Omaha language: revitalization .....	121
APPENDIX.....	124
REFERENCES.....	126

## **ACKNOWLEDGEMENTS**

First, I would like to thank the Omaha people who will help me with my study and consent to work with me. I also thank them for their teaching the small amount of Omaha we had time for during my field work. Their input was indispensable and their patience impressive. I would like to thank my thesis committee chair Lyle Campbell, and committee members Rachel Hayes-Harb and Ardis Eschenberg for their extensive help with my drafts. I also appreciate Ardis's kindness in opening her home to me during part of my time in Nebraska. I am also extremely grateful for the advice and direction of Bob Rankin and John Koontz, Siouan linguists who volunteered their time to help answer my many questions about Omaha linguistics, as well as supplying numerous unpublished articles on the Omaha language. I thank Naomi Palosaari for lending me her recording equipment, time, and advice.

## I. BACKGROUND

### Documenting the world's speech sounds

Some contention exists about the identities of certain of the world's speech sounds. The linguist's perception and the manner of documentation are two factors to consider. The IPA has attempted to reflect all possible speech sounds, but is controversial to some phonologists and phoneticians (Ladefoged 1990). The International Phonetic Alphabet (IPA) is intended to be a set of symbols for representing all the possible sounds of the world's languages (Ladefoged 1990:550), and is a compilation of all phonemic sounds documented across languages. Furthermore, the 1990 report on the status of the IPA specifies that "possible sounds . . . means phonologically contrastive elements within a language, stating that "the sounds that are represented are primarily those that distinguish one word from another." (Ladefoged 1990: 550). Its symbols represent "distinctive feature combinations" but "the Association nowhere suggests that these categories are either the sole or the best set of features for linguistic purposes" (Ladefoged 1990:550). And on the other hand, some view any observation on phonological inventories as "an essentially impressionistic account based on a linguist's experience of a number of languages" (Maddieson 1984:1).

Contention about the acoustic properties of phones, and ways to document them, exists on an individual language basis too, and is the focus of this paper. In Omaha, a three way phonemic stop contrast has been documented, which exists at bilabial, dental,

and velar places of articulation. The exact categories are purported to be voiced, (b, d, g), voiceless unaspirated (p, t, k), and voiceless aspirated (p<sup>h</sup>, t<sup>h</sup>, k<sup>h</sup>). The purpose of this study is not to establish the phonemic status of these categories, but to measure voice onset time (VOT) for words containing exemplars of the attested contrast.

Phonetic level transcription may not always be accurate without the aid of instrumentation beyond the linguist's own training. Even linguists with good training and good perception filter what they hear through their native phonologies and accents (Bowerman 2008). The task of documentation is of course vital to the preservation and revitalization of endangered languages, but linguists have not always had and still may not always have access to the most precise ways of measuring acoustic values in the field, and cannot always determine minimal pairs to establish phonemes. Thus, when establishing phones in a language, whether they are phonemic or allophonic, it is important to include an acoustic analysis.

### Second language acquisition, speech perception, and the field linguist

As mentioned previously, one motivation for the current study is the question of accuracy during phonetic documentation and transcription. Because it is possible that even trained linguists may not always accurately perceive speech sounds in the language being documented, some transcriptions may reflect perceptual errors, even when utmost care is exercised to avoid these. As *Linguistic Fieldwork* notes, "[the field linguist] should be familiar with basic tools for phonetic analysis, such as how to read a spectrogram. . . a well trained ear can pick up a lot of information, but . . . it is easy to be misled" (Bowerman



2008:68). According to Ladefoged (2003), via Riney et al. (2007:94), ‘any description of the phonetic structures of a language should include an account of the VOTø

Thus, the following background summarizes some research on L2 (second language) speech perception acquisition, and how this applies to adultsø even trained linguistsø listening to a second or foreign language.

A few issues influence second language speech perception by adult learnersø including field linguists. First, certain parts of the listenerø first language (L1) can speed L2 perception acquisition, and other parts can get in the way of acquisition. On one hand, knowledge of L1 sound systems can hinder perception of L2 sounds. (Davidson et al. 2007; Broselow 1987) ‘Interferenceø is defined as L1 traits carrying over to another language that one speaks (Campbell and Mixco 2007). On the other hand, negative transfer such as interference is not responsible for all learner errors. It is widely known that there is a ‘*subset* [italics mine] of errors which are caused by transfer [from the L1]ø (Broselow 1987:262). Of course, ideally trained phoneticians can more accurately discriminate between non-native speech sounds than untrained non-native speakers of languages (Abramson and Tingsabadh 1999).

Positive transfer can also occur. There is evidence suggesting that even noncontrastive phones present in the L1 can speed the perception of the same speech sounds present at the phonemic level in L2, if information is accessed as a surface representation (Curtin et al. 1998). Also, the level of representation accessed by the listener can affect perception (Hayes-Harb 2007). In other words, ‘it appears that L1 surface features can be lexicalized in L2 acquisitionø(Curtin et al. 1998:399). So, the

linguists' perception of non-native phones can depend on whether she accesses the lexical level of items she hears. However, even in light of accounts of both positive and negative transfer between the L1 and the L2, it is reasonable to imagine that Dorsey and other early descriptive linguists may have made mistakes in their phonetic level transcriptions due to their perception of speech sounds, thus justifying this study's reexamination of the attested stop series in Omaha.

### VOT across languages

Existing research on VOT does include quantified acoustic information for the categories of stops which have been documented across languages. Lisker and Abramson (1964) present data indicating that ranges and averages of VOT can serve to distinguish manner categories, especially voiced, voiceless unaspirated, and voiceless aspirated. In the current study, VOT ranges and averages in Omaha were determined and compared to those documented by Lisker and Abramson (1964) in 11 languages with two-, three-, and four-way stop contrasts. The phonetic categories of Omaha contrasts were based on VOT similarity to Lisker and Abramson's VOT findings for manner categories. The reason it is compelling to compare VOT values of other languages with voiced, voiceless unaspirated, and voiceless aspirated phones is that there are additional phonetic categories along the VOT continuum proposed by some linguists.

For example, Riney et al. (2007) suggest that Japanese contains a stop category whose VOT range falls somewhere between those of the prototypical voiceless unaspirated stop, and voiceless aspirated stop. Lisker and Abramson (1964) also report

this category for Korean, calling it ‘slightly aspirated’ Lisker and Abramson (1964) also propose another series of phones that do not seem to fit cleanly into voiced, voiceless unaspirated, and voiceless aspirated categories: that of voiced aspirated stops (e.g.  $b^h$ ,  $d^h$ ,  $g^h$ ). The mean VOTs for this series are slightly shorter than those of traditional voiced stop VOTs. Cho and Ladefoged (1999) mention voiceless velars whose VOTs also seem to fall between what Lisker and Abramson termed ‘short lag’ (unaspirated) and ‘long lag’ (aspirated) stops. While Cho and Ladefoged (1999:223) do not consider VOT findings ‘to lend themselves to statistical clumping procedures’ they do suggest that there are four phonetic categories (they do not include voiced, negative VOTs in their assessment). These are unaspirated (~30ms), slightly aspirated (~50ms), aspirated (~90ms), and a final category for highly aspirated stops as in Tlingit and Navajo. Omaha results are compared to those of Lisker and Abramson in order to judge whether Omaha stops fall into the most common categories (voiced, voiceless unaspirated, and voiceless aspirated).

#### Omaha language: documentation

Most of the original documentation of Omaha was done by James Owen Dorsey in the late 1800s. He also spent 1878-1880 on the Omaha reservation by which time he was able to communicate with the Indians without an interpreter. Dorsey’s work includes large volumes of transcriptions of Omaha written correspondence and stories, glossed in a pre-IPA orthography. His other linguistic contributions to the documentation include a study of Omaha onomatopoeia. Additional documents were unpublished at the time of his death in 1895 but many manuscripts are more anthropological than linguistic in nature, \

which is also true of other researchers' work on Omaha. Dorsey is responsible for 'Dorsey's Law', which is 'a rule inserting epenthetic copy-vowels between the members of certain consonant clusters in various Siouan languages (Hale 1985:427). Dorsey's work with Omaha appears to comprise predominantly descriptive linguistics and anthropology.

Other linguists and anthropologists who have worked specifically with Omaha include Alice Cunningham Fletcher, John Koontz, Francis LaFlèche, Bob Rankin, Catherine Rudin, Hans Wolff and others. Koontz 1984 (via Eschenberg 2005:4) 'provides an in depth sketch of Omaha-Ponca within a modern linguistic framework. He concentrates mainly on phonology and morphology, extensively discussing how various morphemes combine and what their phonological output is

Inconsistency is one problem with the Dorsey Dhegiha transcriptions. First, because Dorsey frequently confused them, voiceless stops written upright may represent either unaspirated voiceless consonants or geminates. Eschenberg and Rankin were consulted to minimize accidental inclusion of geminates in the present study. Also, Dorsey sometimes misrepresented stop consonants as aspirates (Rankin 2008). Next, Dorsey often used a *brevet* over vowels, the meaning of which is unknown. Rankin suggests retranscribing such vowels without diacritics (2008). Finally, Dorsey often transcribed *ʔkhi* (velar aspirate) when he should have written *ʔkki* (velar unaspirated). In his 2008 retranscription of Dorsey's original manuscripts, Rankin attempts to correct these errors. The current study uses many words from this retranscription, but the exact phonemic category attested in some words will be called into question based on results

from the VOT analysis. Some words have been transcribed in multiple ways throughout the manuscripts, and thus should be reexamined.

Omaha phonology has also been documented by Wolff (1950), whose phoneme inventory includes only the following stops: voiceless unaspirated stops are not mentioned:

/p/, /b/, /t/, /d/, /k/, and /g/.

Finally, Eschenberg (2005) composed a table of the Omaha phonemic index (Table 1).

#### Omaha language: revitalization

Revitalization efforts of Omaha are described in detail in the discussion (below). As stated above, L2 acquisition is influenced by the L1. In the case of Omaha language learners: most of whom are Omaha tribe members but L1 English speakers: it is necessary to learn for Omaha a three way stop contrast that does not exist phonemically in the L1. If we assume the three way contrast is phonemic, Omaha learners access L1 (English) surface level phones (unaspirated stops) which are present at the phonological level in the L2 (Omaha). For example, ɤpotø /pat/ and ɤspotø /spat/ are actually pronounced [p<sup>h</sup>at] and [spat]. Thus at the phonetic level we see two versions of /p/ in English: aspirated (/pat/) and unaspirated (/spat/). Compared to learners whose L1 background does not include both unaspirated and aspirated phones, these learners are at an advantage (Curtin et al. 1998). However, they are still being asked to perceive and produce meaningful language containing purported phonemes that do not exist in

English. It will be useful for language instructors to be aware of and focus on the need to produce these speech sounds, which are not phonemic in English. This will be the case whether the phonetic identity of a stop reflects the hypothesis (e.g., ʔunaspirated) or does not (e.g., ʔslightly aspirated). Being aware of the phonetic identity of the stops will permit refinement of the pronunciation curriculum accordingly.

### Overview

This study measures in milliseconds the VOT<sub>0</sub> defined as the time between the plosive burst and the onset of vocal fold vibration<sub>0</sub> of bilabial, dental, and velar stops (claimed to be voiced, voiceless unaspirated, and voiceless aspirated) in Omaha. This sheds light on the phonetic identity of these stops, which have not previously been analyzed acoustically. This is important specifically because there is evidence to support that adult L2 language learners are less able to discriminate sounds if they do not either contrast phonemically or surface phonetically in the L1 (Curtin et al. 1998). Thus, it is possible that even well trained linguistic field workers miss phonetic nuances because they are unable to perceive them as the native speakers can. Because it is known that there are more than three possible categories of oral stop (see Lisker and Abramson 1964; Cho and Ladefoged 1999, Riney et al. 2007), it is possible that the purported Omaha stop categories have mean VOTs that align more closely with another category than that in which they have been placed.

In the discussion section, an important application is also addressed: revitalization efforts for the Omaha language. It is intended that the results from this study can be

applied to the revitalization efforts for Omaha, especially in regards to traditional pronunciation.

Table 1 Consonant inventory of Omaha (from Eschenberg 2005)

	Bilabial	Interdental	Dental	Alveopalatal	Palatal	Velar	Glottal
Stop	b p p <sup>h</sup>		d t t <sup>h</sup>			g k k <sup>h</sup>	ʔ
Ejective	p'		t' s'	ʃ'		k'	
Affricate				tʃ tʃ <sup>h</sup>			
Fricative		ð	z s	ʒ ʃ		x ɣ	
Glide	w				(y) <sup>2</sup>		
Nasal	m		n				



## II. HYPOTHESIS

Regarding the individual stimulus items examined in the present study, it is hypothesized that because of negative transfer from linguists' L1 (Davidson et al. 2007; Broselow 1987), occasional transcription errors from past Omaha documentation will be present in the stimulus items used for this experiment, which will come to light during the VOT analysis, e.g. a phone transcribed as an unaspirated voiceless stop, when recorded, may have a VOT that falls within the ranges typically documented for aspirated stops. This will become noticeable if outlying VOT values during the analysis of a certain phone can be linked to a specific word or group of words. Cross linguistic analysis shows with all places of articulation that voiceless aspirates are much less likely to be present in languages than either plain voiceless or voiced stops (Maddieson 1984:205-215). Based on this, it is predicted that VOT analysis will reveal parallel proportions across the stop series.

Regarding the overall categorization of the three way stop series, it is hypothesized that data points accumulate in the typical three categories: voiced, voiceless unaspirated, and aspirated, because these are the most common cross linguistically (Maddieson 1984). All other series occur less frequently. This outcome is expected in spite of possible transcription and perception errors of certain specific words by linguists who have documented Omaha in the past.

### III. METHODS

#### Stimuli

Stimuli consist of single words in Omaha elicited in the context of a carrier phrase in Omaha. A total of thirty-six tokens was elicited three times each. Words were chosen from transcriptions by Dorsey from c. 1890, with help from Ardis Eschenberg, a linguist who has worked on Omaha and lived 10 years on the Omaha reservation. When Dorsey's transcriptions were superseded by Eschenberg's suggestions, it was due to known transcription errors in the Dorsey text or to the antiquated nature of some qualifying words in the manuscripts. Word length is not controlled because preference was given to plausibility of elicitation, although when possible, shorter words were selected. The limitation of this approach is that other syllables in target words potentially could affect the value of the target stop consonant.

For each place of articulation, the stop series was elicited in syllable onset position preceding each of the attested vowels in Omaha except /o/. Because stress is known to influence VOT variably across languages (Cho and McQueen 2005), only target syllables purported to be unstressed were included (stressed syllables do exist in other syllables in all words). In the final word list (written in Fletcher-LaFlèche orthography), 12 phrases contain bilabial stops, 12 contain dental stops, and 12 contain velar stops. Of these, three each contain /a/, /e/, /i/, and /u/. There was no data found for

stops preceding /o/ among bilabial and dental stops. Typically, /o/ is found only in a suffix denoting a male speaker and in the greeting used by male speakers (Eschenberg, personal communication). However, three words were found where velar stops preceded /o/.

The breakdown of stimulus items is shown in Table 2; contextualized syllables are shown in Table 3.

As mentioned, this study does not control for vowel length, which is phonemic (Rankin, Eschenberg, personal communication). Also, nasalized vowels have been deliberately omitted from all target syllables. Although tokens regularly include nasalized vowels, these are never part of the syllable containing the stop consonant under consideration. Thus, there are several variables that are not controlled in the word list: vowel length, nasal influence on target syllable, and word position.

The questionnaire was constructed by randomizing the entire list of three repetitions of each of the 36 tokens. An online random sequence generator will be used to do this. To avoid list intonation associated with the first and last items on a page, two fillers will be included at the beginning and end of each page. Thus, these items are in addition to the 108 items that will actually be considered in the study.

All tokens were presented in a contextualized carrier phrase. Each word's carrier phrase is semantically appropriate. Although this diminishes the consistency of the phonological environment compared to using a single carrier phrase for all words, it greatly improves chances of eliciting each word, especially in the case of words that are infrequently used and difficult for speakers to recall. Therefore, in addition to the

noncontrolled variables listed above it is necessary to note that carrier phrases vary among token words and are not controlled.

### Subjects

Four fluent speakers and two semifluent speakers of Omaha were recorded. Of the two semifluent speakers, one had learned some Omaha successfully in an L2 classroom setting as well as having limited familial exposure to the language. In order to maximize the sample size additional criteria such as speech and hearing impairments were not designated. Very few Omaha speakers remain today, making a larger sample size nearly impossible to obtain.

### Data collection

An M-Audio Microtrack 24/96 recorder was used in all elicitation, and the recordings saved to a PC hard drive and an external hard drive. Participants wishing to obtain a copy of their recording will receive a DVD. A behind the head microphone (Listen brand LA-278) was used with the recorder. All files were mono recorded in .wav format, with a sample rate of 48 kHz , and 24 bit resolution. iTunes was used to import the files. The target syllables containing examples of the stop series was spliced from the carrier phrases using Praat and saved individually. After isolating the target syllables in the recordings, Praat was used to measure VOT.

During the elicitation, an older native speaker assisted two of the speakers by clarifying in English word meanings and contextualized words for speakers who had not

used Omaha in many years. When assistance was necessary, she provided it with the aid of a written copy of the questionnaire. All orthography was in the Fletcher-LaFlèche style, which is more familiar to speakers than IPA.

Table 2

Target syllables (stimulus items), which were embedded in carrier phrases.

Bilabial	Dental	Velar	Bilabial	Dental	Velar
/ba/	/da/	/ga/	/be/	/de/	/ge/
/pa/	/ta/	/ka/	/pe/	/te/	/ke/
/pha/	/tha/	/kha/	/phe/	/the/	/khe/
/bi/	/di/	/gi/	/bu/	/du/	/gu/
/pi/	/ti/	/ki/	/pu/	/tu/	/ku/
/phi/	/thi/	/khi/	/phu/	/thu/	/khu/

Table 3

Example of contextualizing words used with target syllables.

<u>Omaha (IPA)</u>	<b>ba té</b>	<b>ába</b>	<b>pahí</b>	<b>nápe</b>	<b>áp<sup>h</sup>ahi</b>	<b>ǫap<sup>h</sup>a</b>
<u>Omaha</u> <u>(Fletcher-</u> <u>LaFlèche)</u>	<b>bashté</b>	<b>ó<sup>N</sup>ba</b>	<b>pahí<sup>N</sup></b>	<b>nó<sup>N</sup>pe</b>	<b>zháp<sup>h</sup>ahi</b>	<b>í<sup>N</sup>thap<sup>h</sup>a</b>
<u>Gloss</u>	strawberry	day	Quill	chokecherries	rosinweed	Corn-crusher
<u>Place of articulation</u>	Bilabial					
<u>Voicing</u>	Voiced		Voiceless unaspirated		Voiceless aspirated	
<u>Stressed vowel</u>	No					
<u>Syllable position</u>	initial	final	initial	final	medial	final

#### **IV. RESULTS/ANALYSIS**

VOT was measured using PRAAT (Boersma and Weenink 2007). For voiceless unaspirated and voiceless aspirated stops (positive VOT), milliseconds were measured between the zero crossing closest to the release burst, and the onset of vocal fold vibration as represented by the zero crossing at first sign of periodic acoustic activity. For voiced stops (negative VOT), measurement was taken from the zero crossing nearest the first low amplitude periodic activity, to the end, marked by the release and onset of high amplitude vocalic wavelengths. For intervocalic negative VOT, closure duration was measured.

Results were calculated in Microsoft Excel and are shown and described below. They are given in order of most general to most specific. Error bars are for one standard deviation of error. Data markers above bars show the number of tokens comprising each average. This is compelling because, due to unsuccessful elicitation of some tokens, bars may represent very different n values, making them more difficult to compare against one another. The charts represent VOT findings expressed in the following ways:

- individual tokens by speaker
- individual speaker averages
- averages by fluency
- averages by fluency and word position
- averages by fluency and stress



- averages by fluency, word position, and stress
- average by fluency, vowel, and place of articulation
- averages by fluency, vowel, place of articulation, word position, and stress

Before the presentation of each cross analysis, a short description of the analysis is given. Peculiarities or unexpected results are noted. After all cross analyses have been presented, a chart of total missing tokens is provided. Following these sections is the discussion section.

#### VOT individual tokens by speaker

The following graphs (Figure 1 to Figure 18) represent individual token millisecond (msec) values as they were elicited from each speaker. Each graph shows only one place of articulation. All tokens are represented individually with one point on the graph, thus, the first three points of any color will represent /a/ syllables. In the bilabials chart, the left-most three green points represent /pha/, while the left-most three red points are /pa/. The next three of each color stand for /phe/ and /pe/ respectively. Because of the setup of this study, only one word has been used for each different syllable. Therefore, all the /pa/ syllables are from just one word, *paɦi<sup>n</sup>ə*. As is visible in some of these charts (e.g. Figure 1, Figure 5, Figure 7, Figure 10, and Figure 15) some overlap occurs between tokens of different stop categories, although category averages are distinctly separate (Figure 19 to Figure 24). Missing points on the graphs represent missing tokens—those repetitions that were not successfully elicited from the speaker.

### Individual speaker averages

These analyses (Figure 19 to Figure 24) show individual speaker averages (msec) across tokens for each place of articulation and manner of stop. All error bars are for one standard deviation. Neither word position nor stress has been isolated here. In all 6 figures it is possible to note the pattern of negative VOTs for voiced stops, relatively short VOTs for voiceless unaspirated stops, and the longest VOTs for voiceless aspirated stops. This pattern prevails (with the exception of Figure 22 bilabials) in spite of lack of conformity by some individual tokens as shown in the section above.

### VOT averages by fluency

These analyses (Figure 25 to Figure 27) represent three groups: all speakers, semifluent speakers, and fluent speakers (by self reported fluency). Neither stress nor word position variables are considered below.

### VOT averages by fluency and word position

This set of analyses (Figure 28 to Figure 33) differs from the previous one in that word position is isolated in addition to fluency level. However, stress differences are not considered here. Word initial syllable results are shown, followed by word medial results. Most target syllables were word medial, making n values for word initial charts very low. In the case of /ph/, /d/, and /th/, there is an n=0 for word initial tokens. Averages in word medial tokens (Figure 31 to Figure 33) are consistent with averages from other analyses: there are negative VOTs for voiced stops, and shorter VOTs for unaspirated than for

aspirated voiceless stops. However, word initial tokens comprise few examples, making it difficult to know if this pattern is consistent for word initial syllables.

#### VOT averages by fluency and stress

The following analysis (Figure 34 to Figure 39) isolates syllable stress, but does not account for word position. Although this study was designed to test only unstressed syllables, certain tokens were found to be stressed during the elicitations. Thus, they are separated below. Stressed-syllable charts reveal low *n* values for stressed syllables compared to those of unstressed syllables. Nevertheless, both sets of data show three way contrast patterns.

#### VOT averages by fluency, word position, and stress

These analyses (Figure 40 to Figure 51) consider a variety of factors simultaneously: all permutations of fluency, word position, and stress are isolated below. The obvious advantage of isolating each of these variables is minimizing confounds in the results that would prevent an independent assessment of each variable's influence on VOT. However, the disadvantage of this analysis for the current study's data is the low *n* values in certain instances. For example, Figure 51 includes only values for /ph/ and /d/.

This prevents any comparison at any place of articulation for any permutation of stop manner combinations.

#### VOT averages by fluency, vowel, and place of articulation

Charts below (Figure 52 to Figure 60) are color coded by place of articulation—blue for bilabials, red for dentals, orange for velars. Each chart separates data by vowel, since stops were elicited preceding each of the four common vowels in Omaha—/a/, /e/, /i/, and /u/. The next analysis accounts for word position and stress, but this analysis does not. The rationale is that viewing averages with higher n values will allow for a clearer comparison although confounds exist.

#### Cross speaker VOT averages by fluency, word position, stress, and vowel

The final analysis (Figure 61 to Figure 94) is the most detailed one included in the study. It eliminates more confounds than any of the preceding analyses. The shortcoming is that there are far fewer tokens in each category because so many are eliminated via cross analysis. For example, there is only one token which fulfills the requirement—word medial stressed velar from a semifluent speaker—

Word initial unstressed stops by vowel:

#### Missing tokens

Figure 95 shows missing tokens. These are tokens which were not successfully elicited but which ideally would have been included in the study.

Speaker A (fluent):

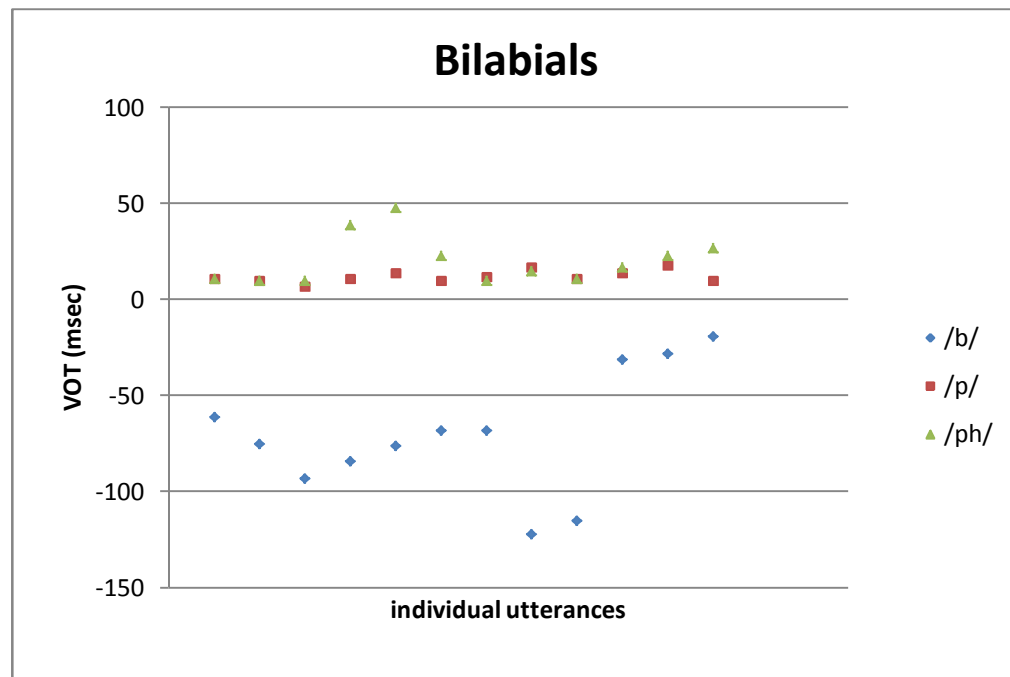


Figure 1.

Talker A's individual voice onset times for each bilabial target consonant individually. The first three tokens represent the target consonant followed by /a/; the next three by /e/; the next three by /i/; the final three by /u/. Missing tokens indicate that the talker did not produce any usable tokens for that target consonant.

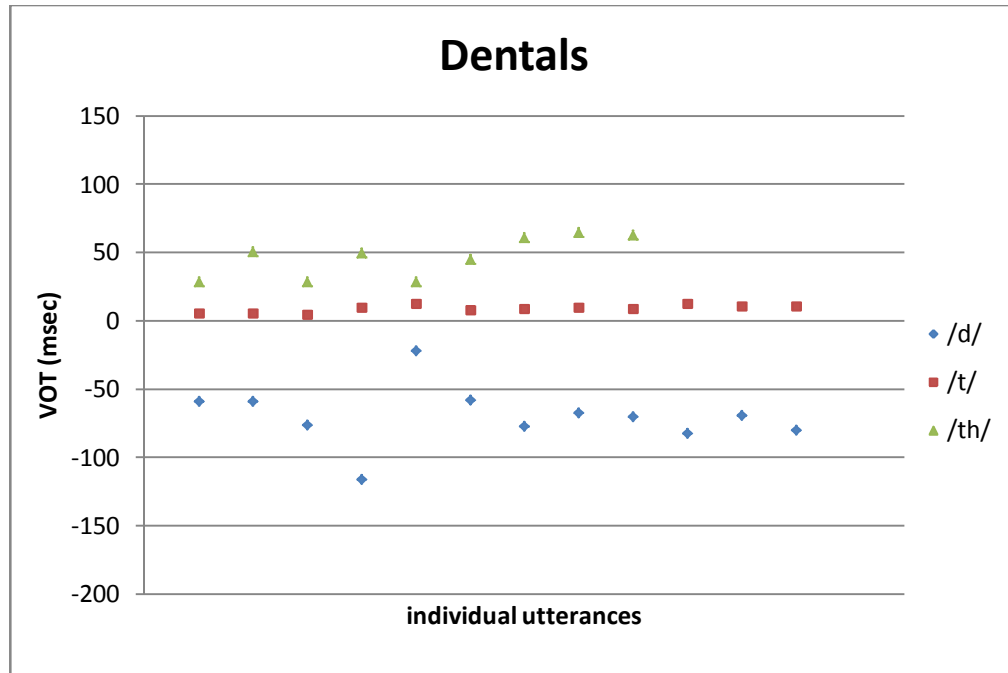


Figure 2.

Talker A's individual voice onset times for each dental target consonant individually. The first three tokens represent the target consonant followed by /a/; the next three by /e/; the next three by /i/; the final three by /u/. Missing tokens indicate that the talker did not produce any usable tokens for that target consonant.

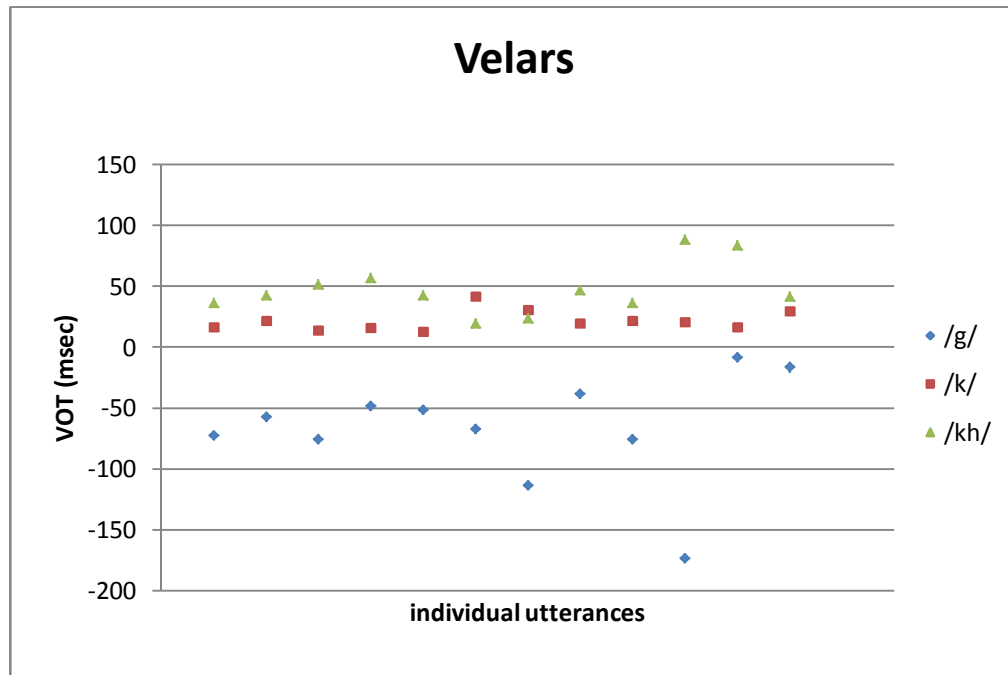


Figure 3.

Talker A's individual voice onset times for each velar target consonant individually. The first three tokens represent the target consonant followed by /a/; the next three by /e/; the next three by /i/; the final three by /u/. Missing tokens indicate that the talker did not produce any usable tokens for that target consonant.

Speaker B (fluent):

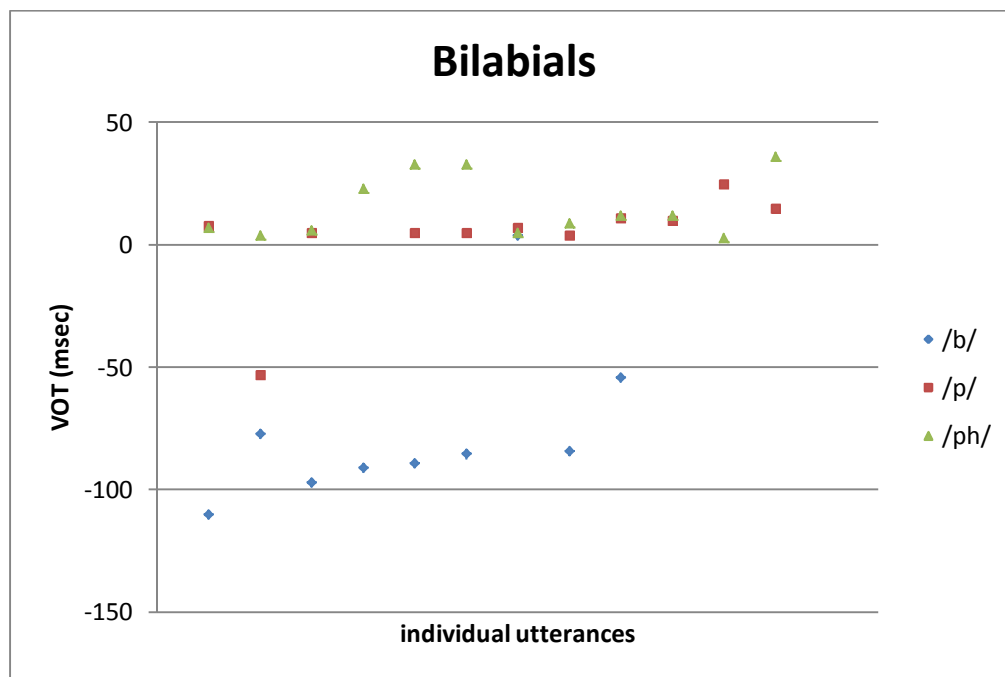


Figure 4.

Talker Bø individual voice onset times for each bilabial target consonant individually. The first three tokens represent the target consonant followed by /a/; the next three by /e/; the next three by /i/; the final three by /u/. Missing tokens indicate that the talker did not produce any usable tokens for that target consonant.



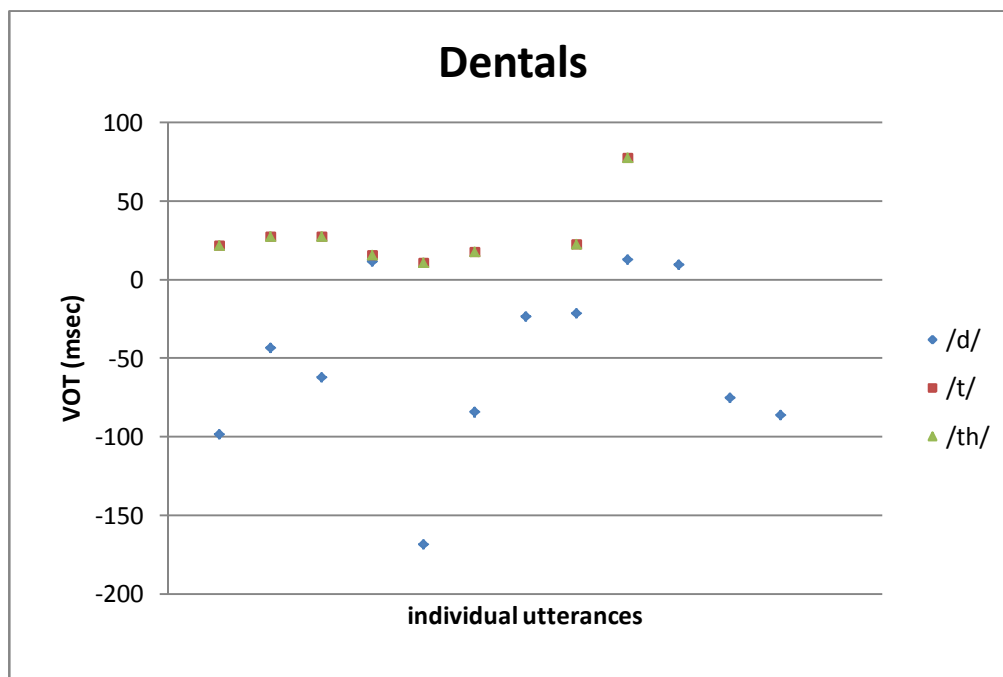


Figure 5.

Talker Bø individual voice onset times for each dental target consonant individually. The first three tokens represent the target consonant followed by /a/; the next three by /e/; the next three by /i/; the final three by /u/. Missing tokens indicate that the talker did not produce any usable tokens for that target consonant.

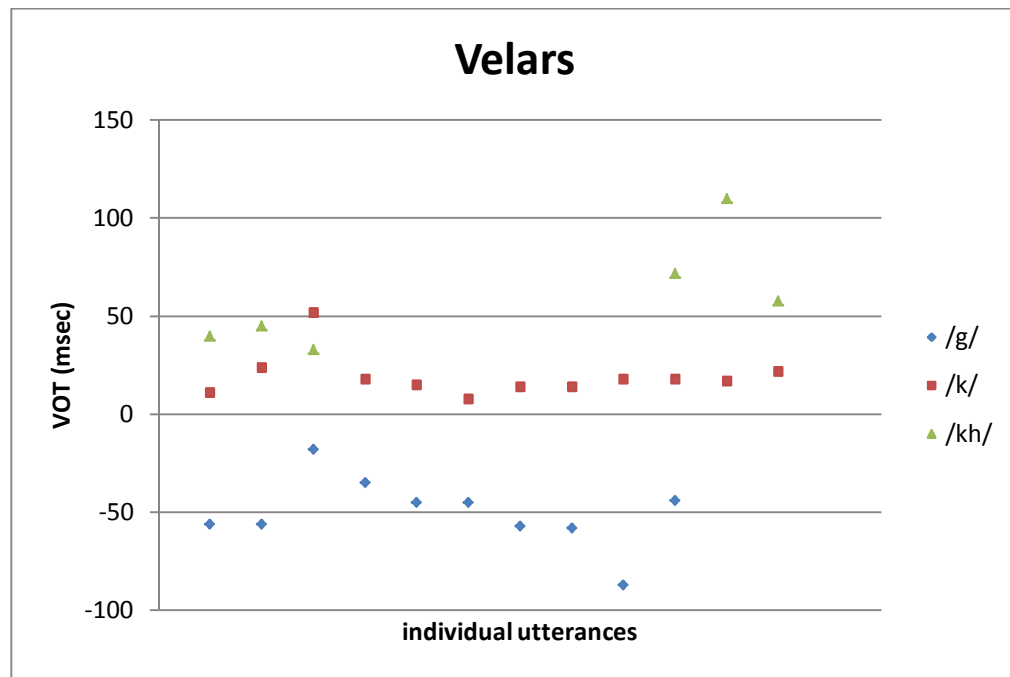


Figure 6.

Talker Bø individual voice onset times for each velar target consonant individually. The first three tokens represent the target consonant followed by /a/; the next three by /e/; the next three by /i/; the final three by /u/. Missing tokens indicate that the talker did not produce any usable tokens for that target consonant.

Speaker C (fluent):

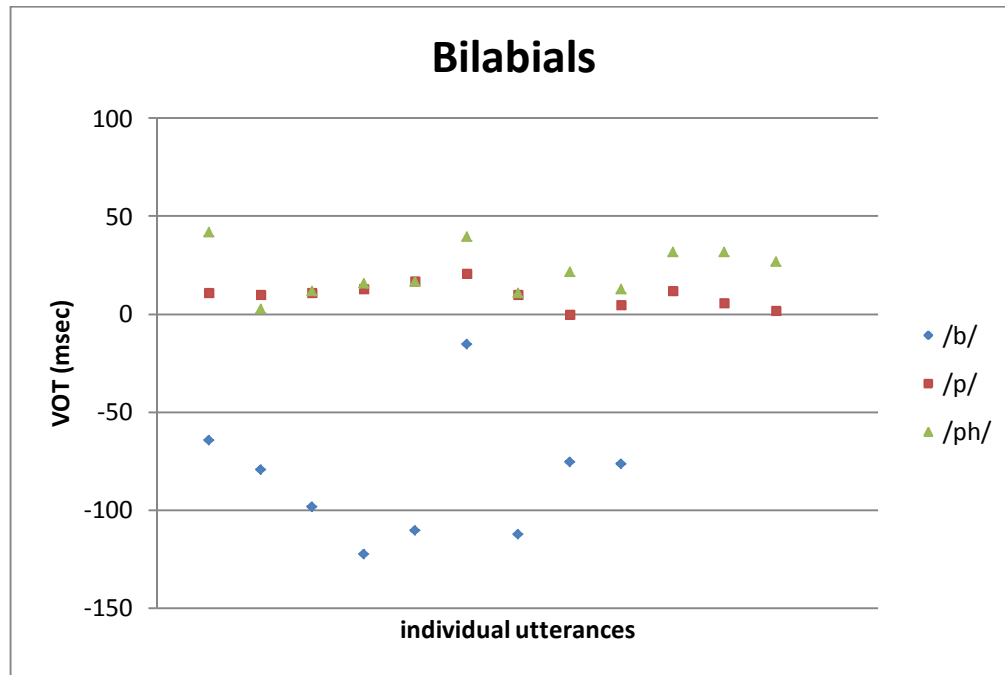


Figure 7.

Talker C's individual voice onset times for each bilabial target consonant individually. The first three tokens represent the target consonant followed by /a/; the next three by /e/; the next three by /i/; the final three by /u/. Missing tokens indicate that the talker did not produce any usable tokens for that target consonant.

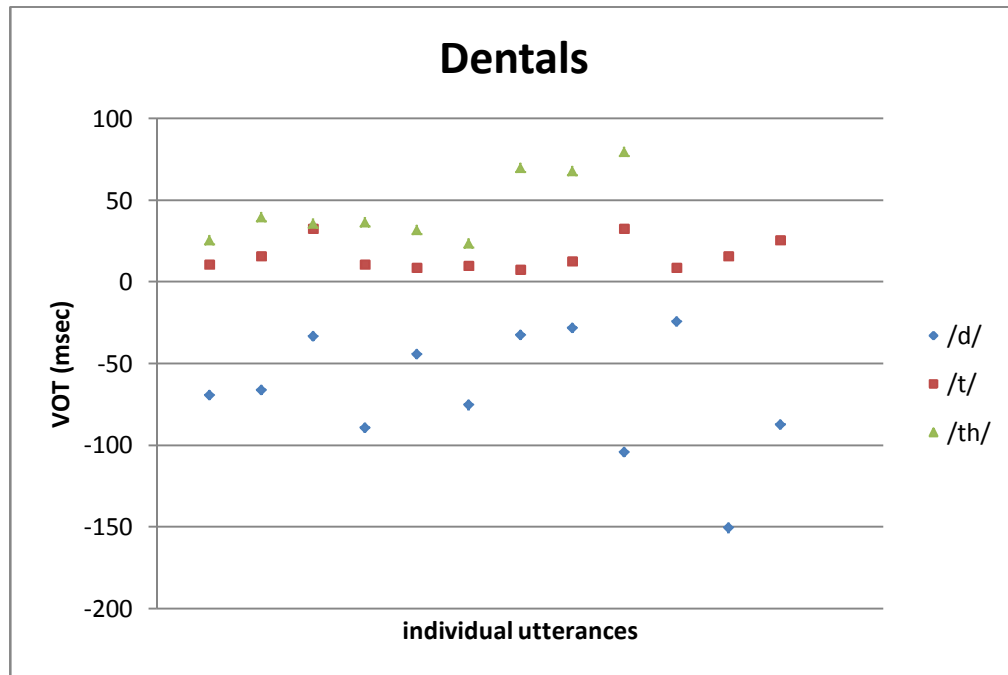


Figure 8.

Talker Cø's individual voice onset times for each dental target consonant individually. The first three tokens represent the target consonant followed by /a/; the next three by /e/; the next three by /i/; the final three by /u/. Missing tokens indicate that the talker did not produce any usable tokens for that target consonant.

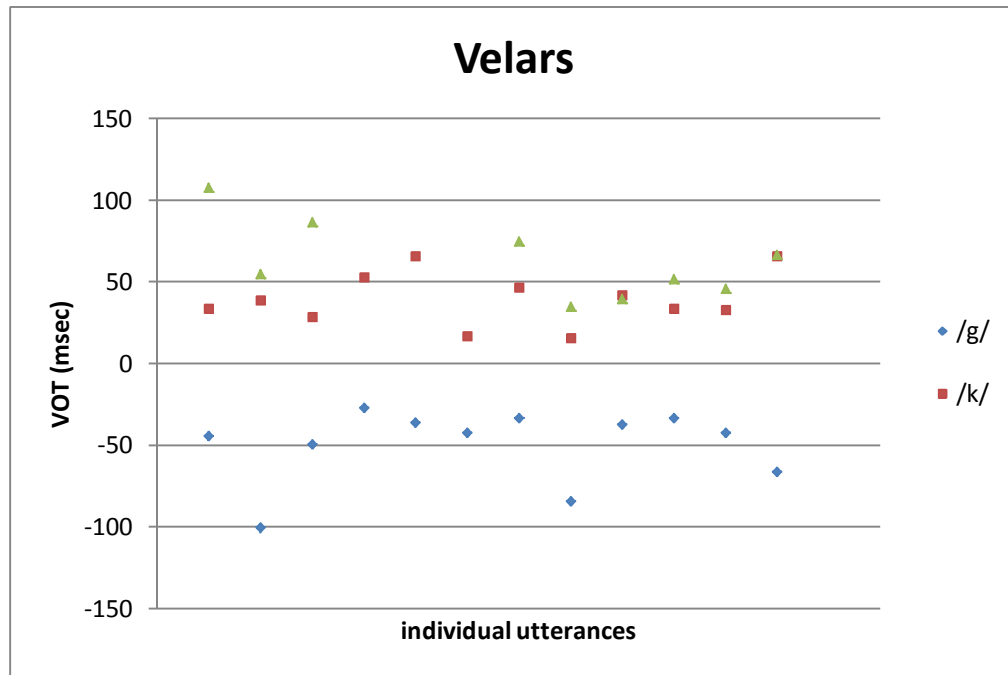


Figure 9.

Talker Cø individual voice onset times for each velar target consonant individually. The first three tokens represent the target consonant followed by /a/; the next three by /e/; the next three by /i/; the final three by /u/. Missing tokens indicate that the talker did not produce any usable tokens for that target consonant.

Speaker D (fluent):

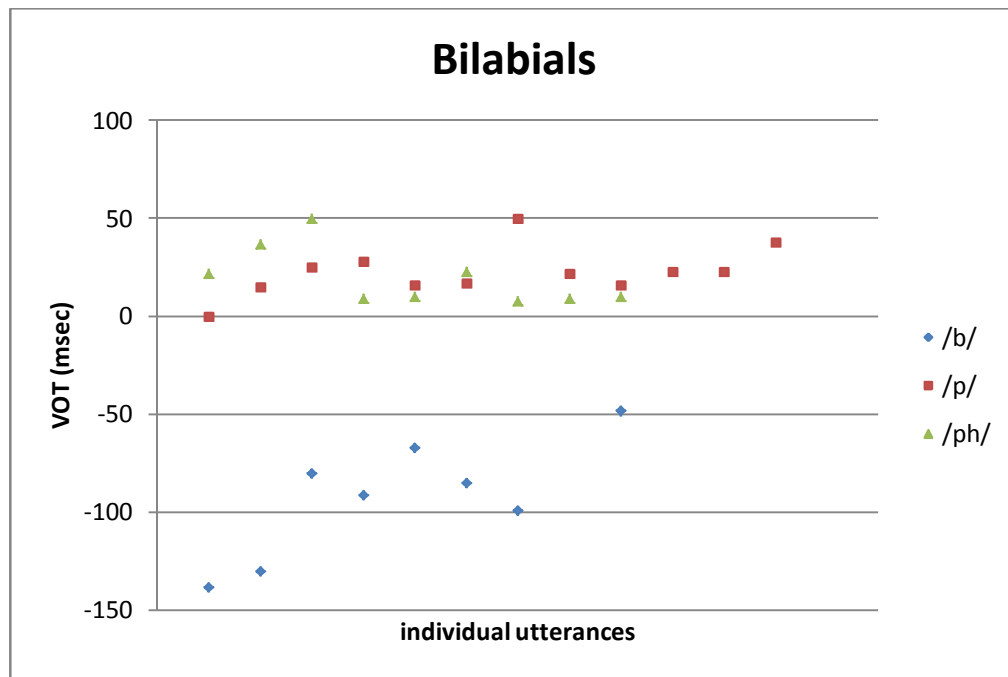


Figure 10.

Talker Dø individual voice onset times for each bilabial target consonant individually. The first three tokens represent the target consonant followed by /a/; the next three by /e/; the next three by /i/; the final three by /u/. Missing tokens indicate that the talker did not produce any usable tokens for that target consonant.

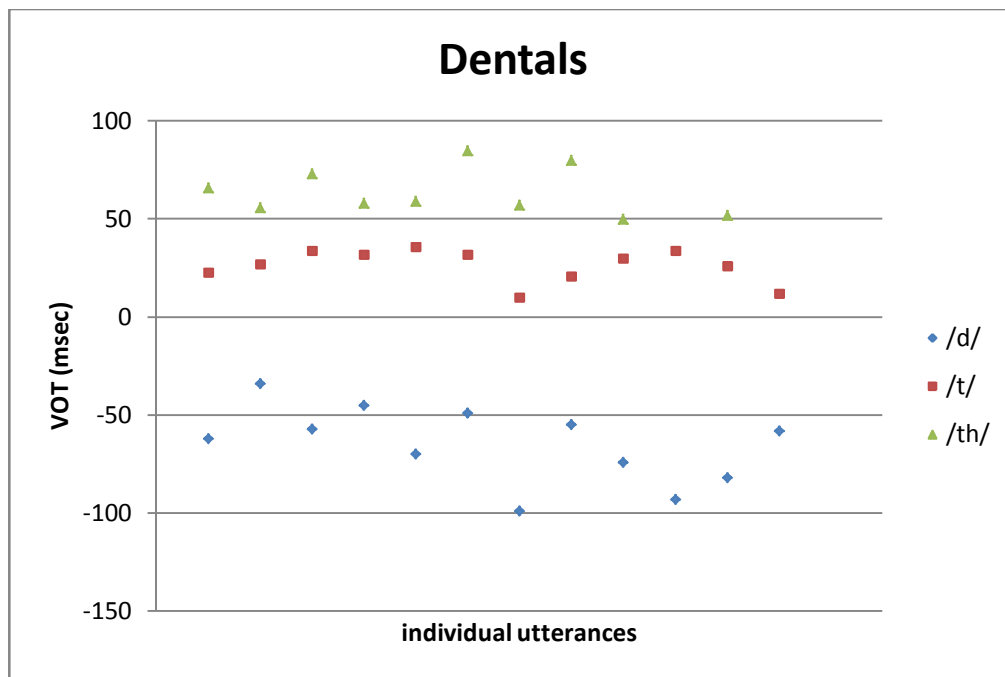


Figure 11.

Talker Dø's individual voice onset times for each dental target consonant individually. The first three tokens represent the target consonant followed by /a/; the next three by /e/; the next three by /i/; the final three by /u/. Missing tokens indicate that the talker did not produce any usable tokens for that target consonant.

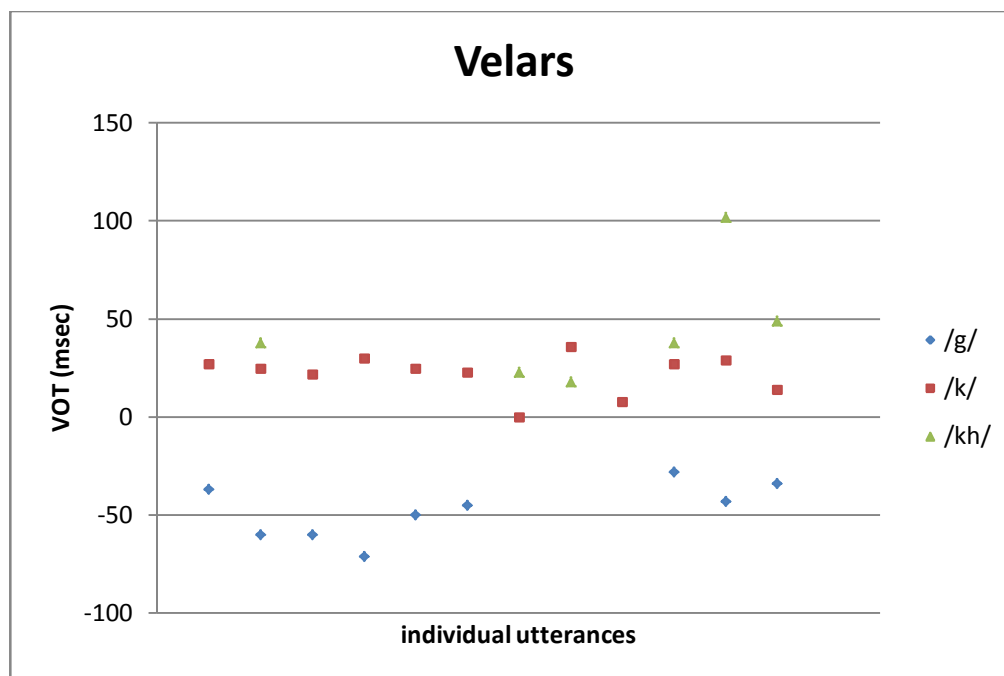


Figure 12.

Talker Dø individual voice onset times for each velar target consonant individually. The first three tokens represent the target consonant followed by /a/; the next three by /e/; the next three by /i/; the final three by /u/. Missing tokens indicate that the talker did not produce any usable tokens for that target consonant.



Speaker E (semifluent):

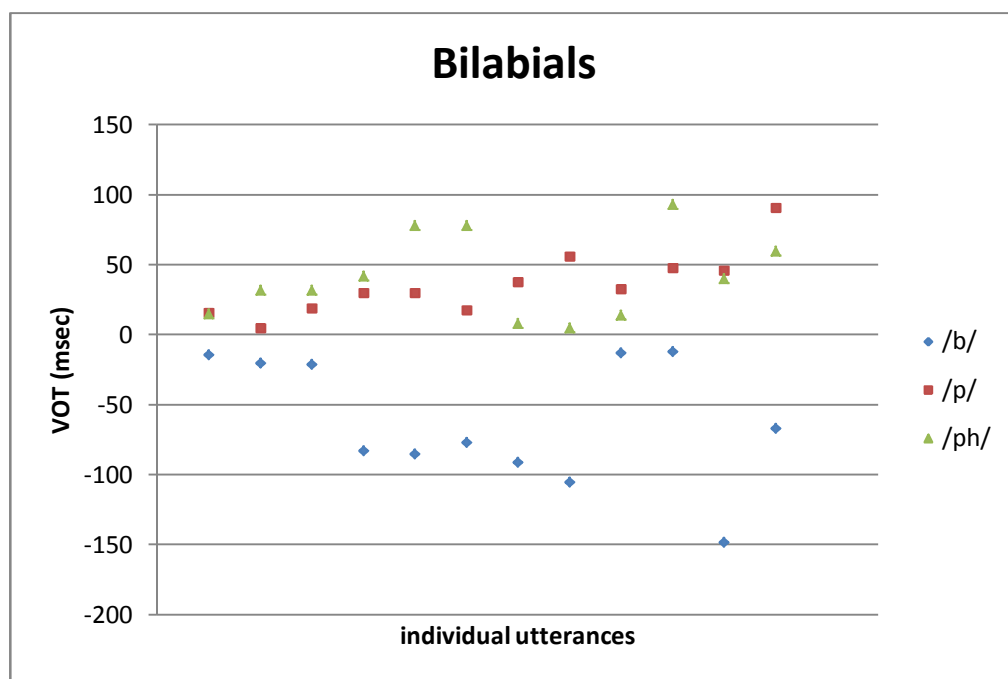


Figure 13.

Talker E's individual voice onset times for each bilabial target consonant individually. The first three tokens represent the target consonant followed by /a/; the next three by /e/; the next three by /i/; the final three by /u/. Missing tokens indicate that the talker did not produce any usable tokens for that target consonant.

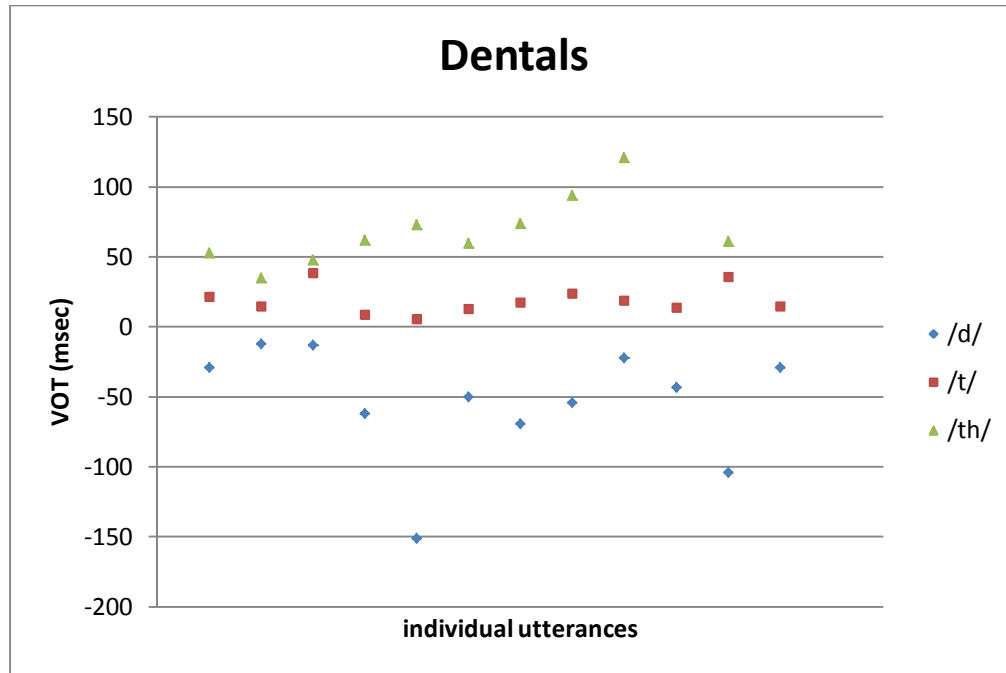


Figure 14.

Talker Eø's individual voice onset times for each dental target consonant individually. The first three tokens represent the target consonant followed by /a/; the next three by /e/; the next three by /i/; the final three by /u/. Missing tokens indicate that the talker did not produce any usable tokens for that target consonant.

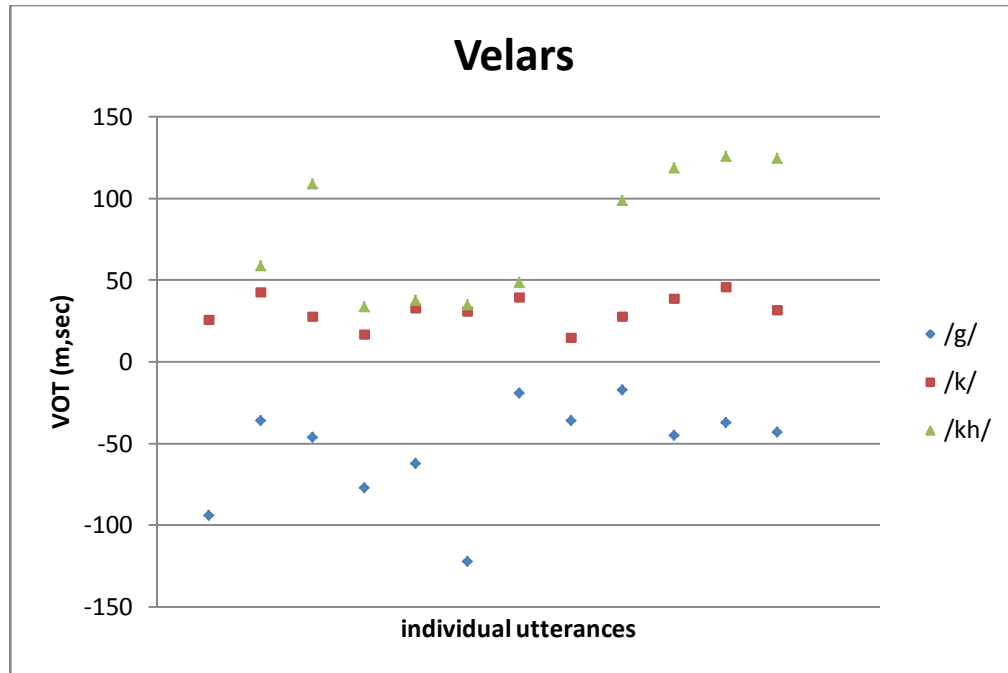


Figure 15.

Talker Eø individual voice onset times for each velar target consonant individually. The first three tokens represent the target consonant followed by /a/; the next three by /e/; the next three by /i/; the final three by /u/. Missing tokens indicate that the talker did not produce any usable tokens for that target consonant.



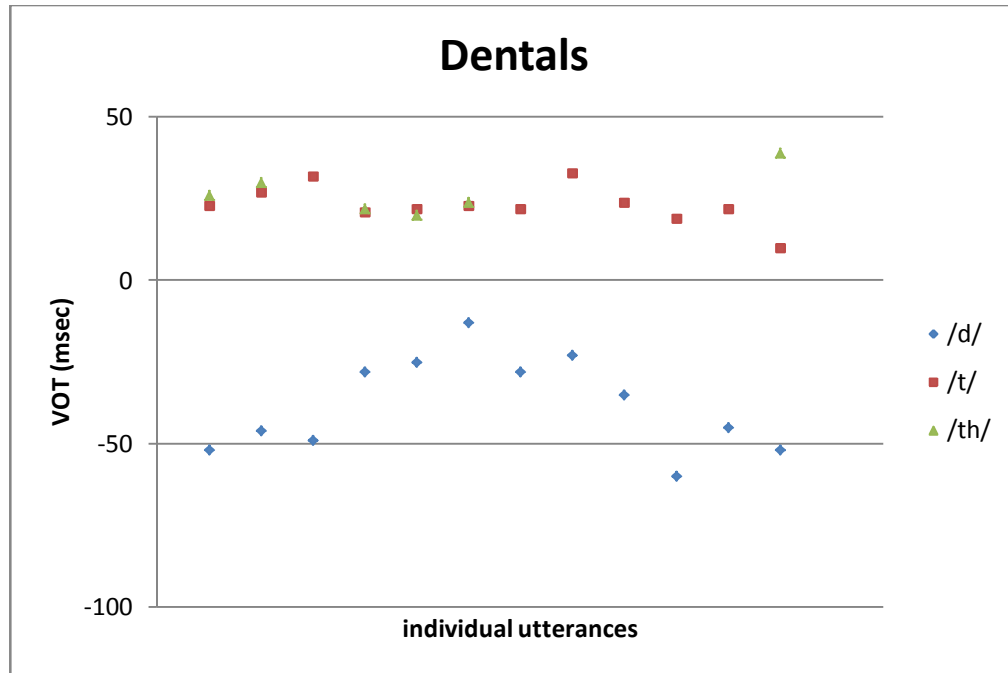


Figure 17.

Talker Fø's individual voice onset times for each dental target consonant individually. The first three tokens represent the target consonant followed by /a/; the next three by /e/; the next three by /i/; the final three by /u/. Missing tokens indicate that the talker did not produce any usable tokens for that target consonant.

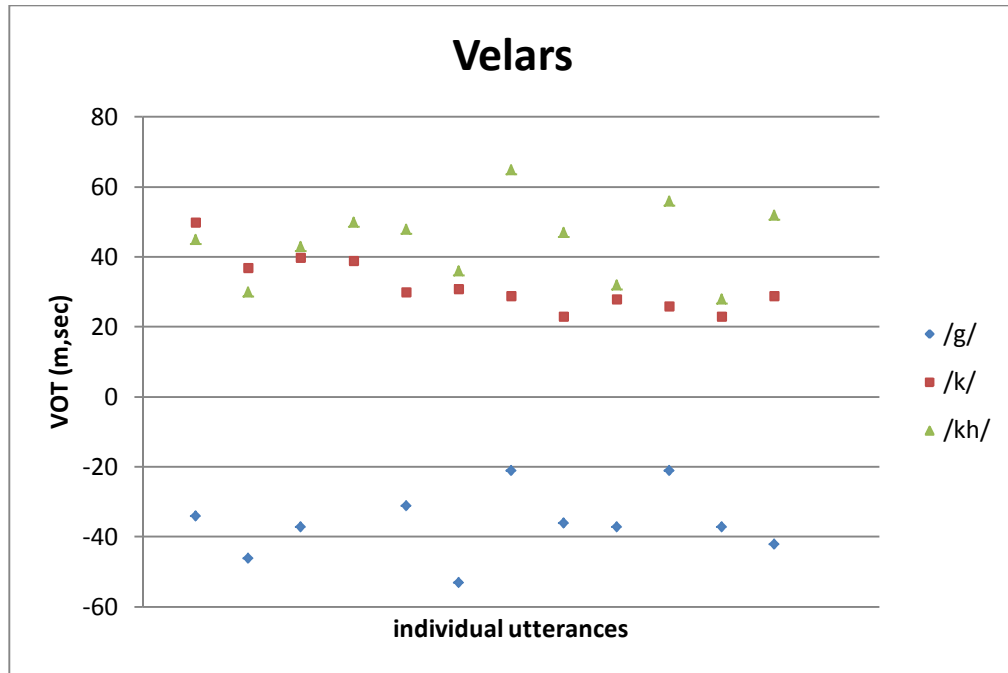


Figure 18.

Talker Fø's individual voice onset times for each velar target consonant individually. The first three tokens represent the target consonant followed by /a/; the next three by /e/; the next three by /i/; the final three by /u/. Missing tokens indicate that the talker did not produce any usable tokens for that target consonant.

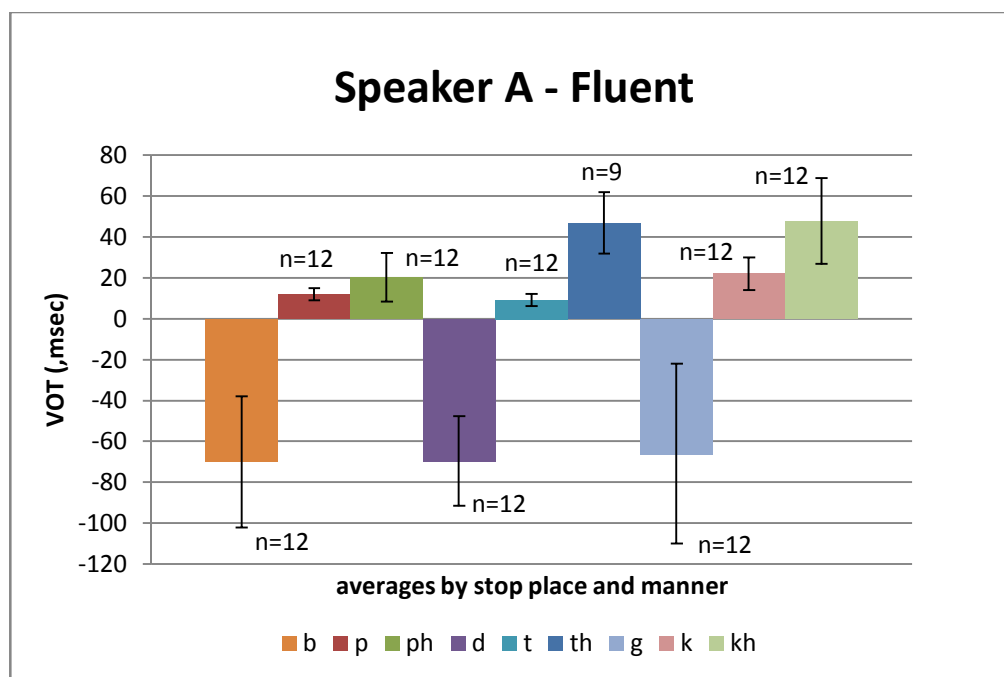


Figure 19.

Talker A's mean voice onset time for each target consonant individually; the number of tokens contributing to each mean is presented above/below the bar.

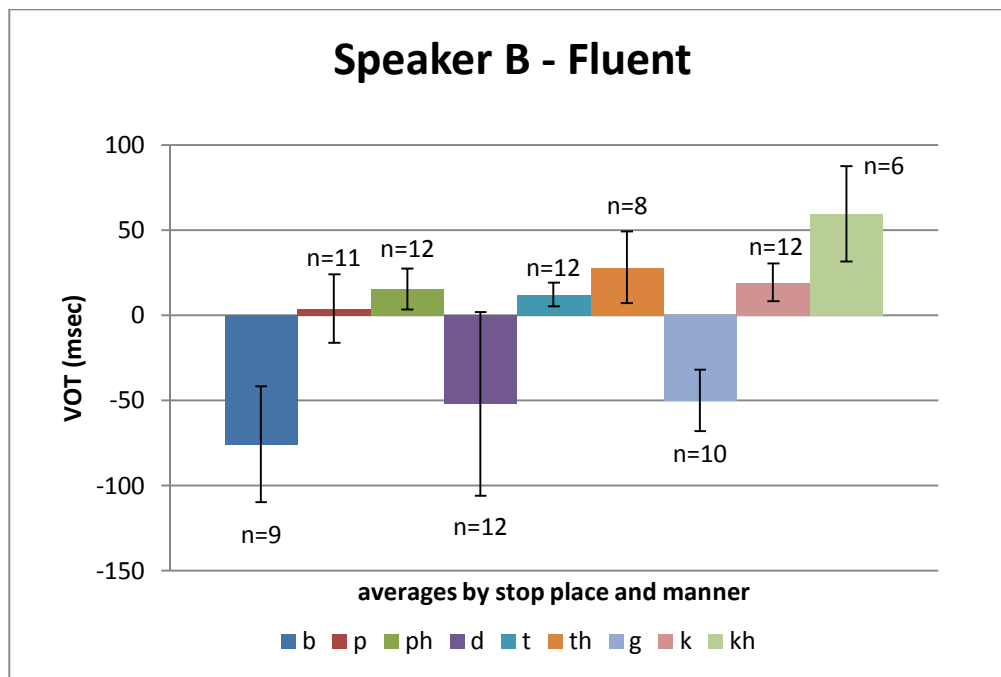


Figure 20.

Talker B's mean voice onset time for each target consonant individually; the number of tokens contributing to each mean is presented above/below the bar.



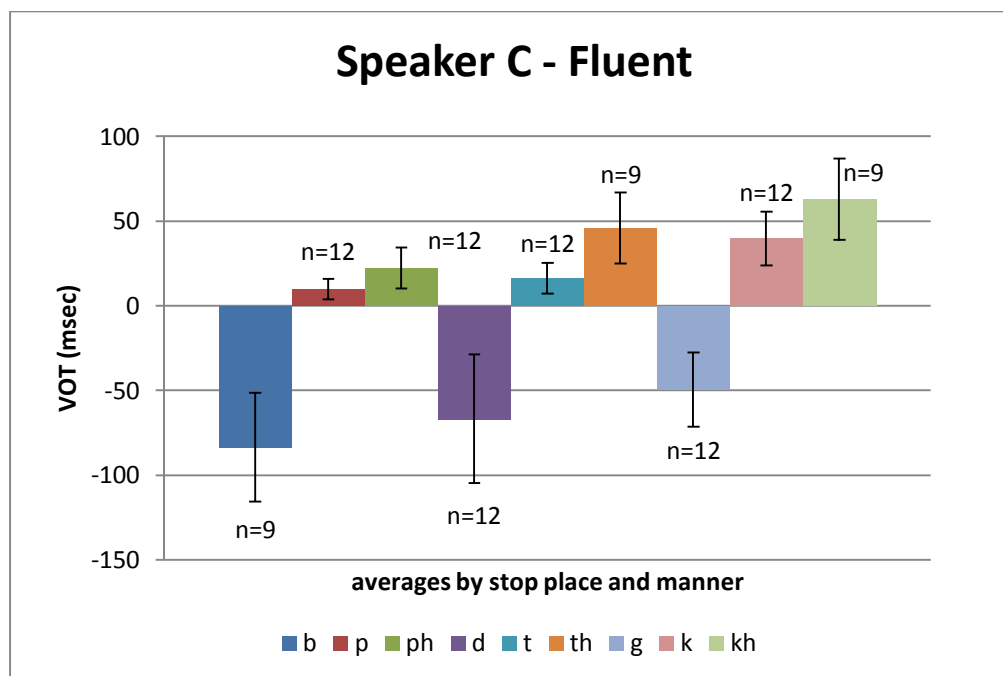


Figure 21.

Talker C's mean voice onset time for each target consonant individually; the number of tokens contributing to each mean is presented above/below the bar.

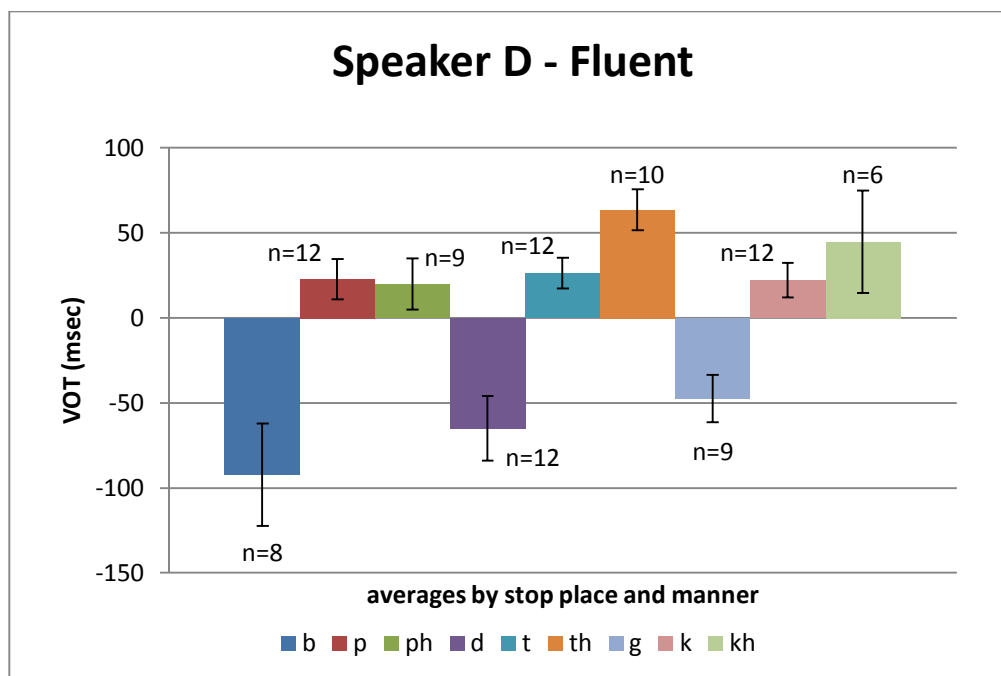


Figure 22.

Talker D's mean voice onset time for each target consonant individually; the number of tokens contributing to each mean is presented above/below the bar.

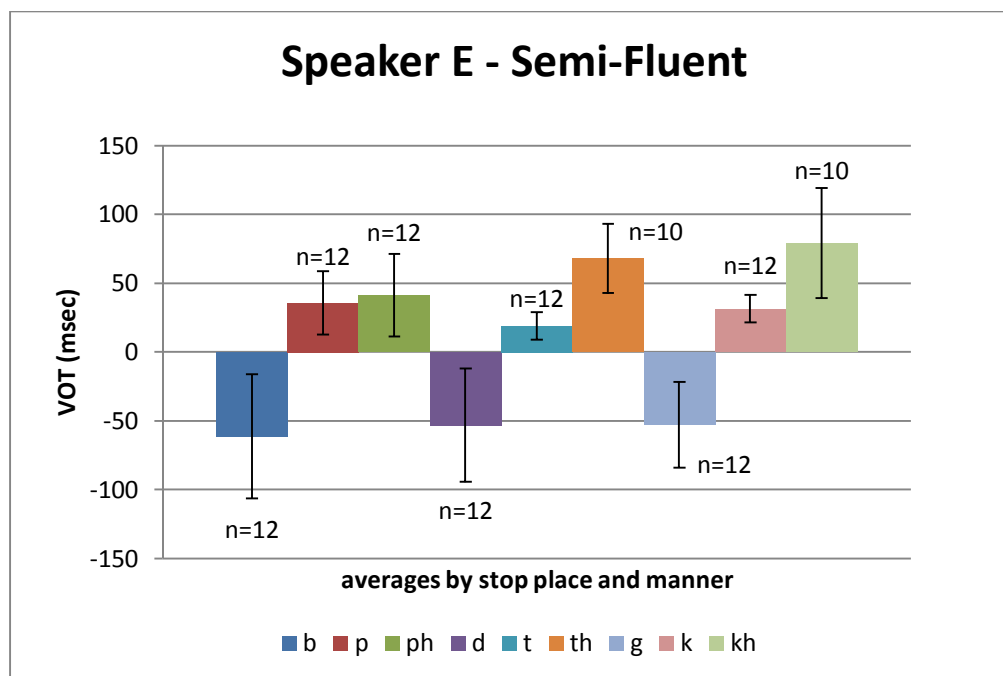


Figure 23.

Talker E's mean voice onset time for each target consonant individually; the number of tokens contributing to each mean is presented above/below the bar.

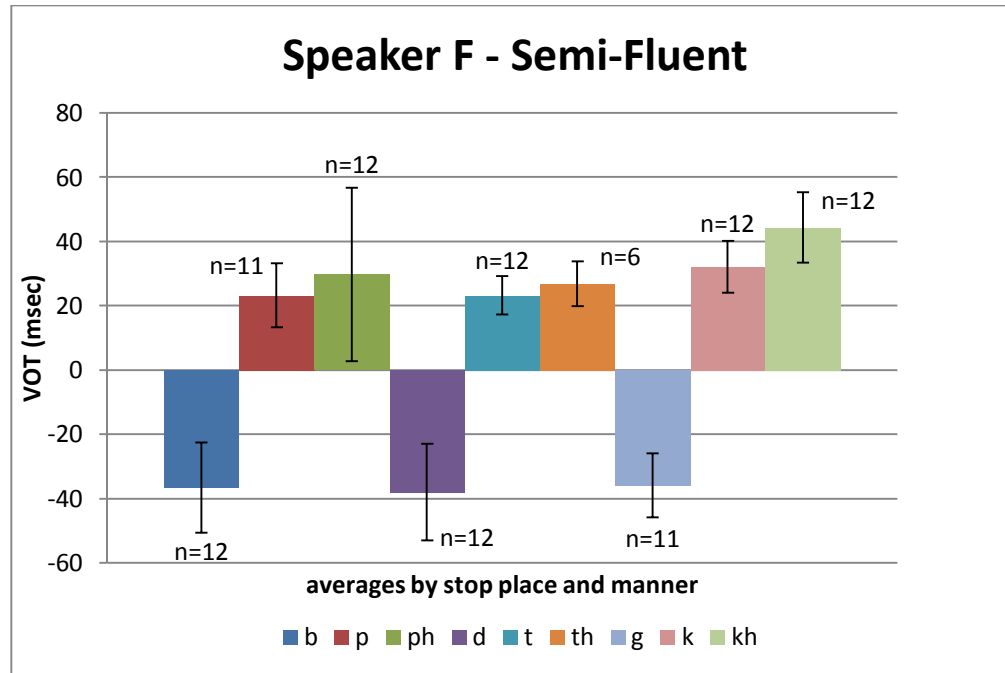


Figure 24.

Talker Fø's mean voice onset time for each target consonant individually; the number of tokens contributing to each mean is presented above/below the bar.

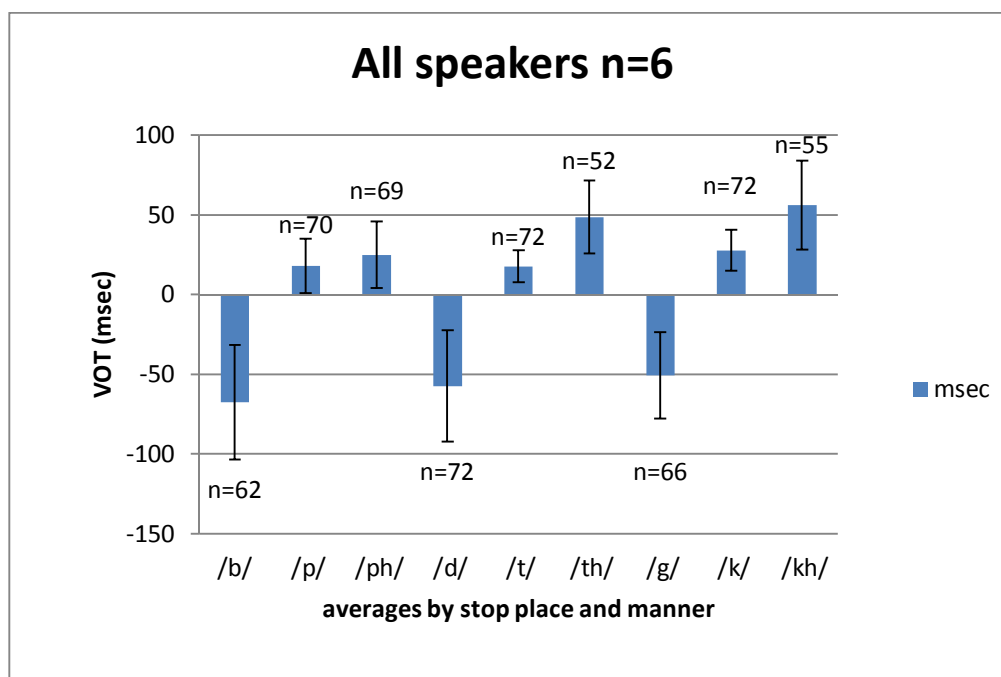


Figure 25.

Talkers A-F (all talkers): mean voice onset time for each target consonant individually; the number of tokens contributing to each mean is presented above/below the bar.

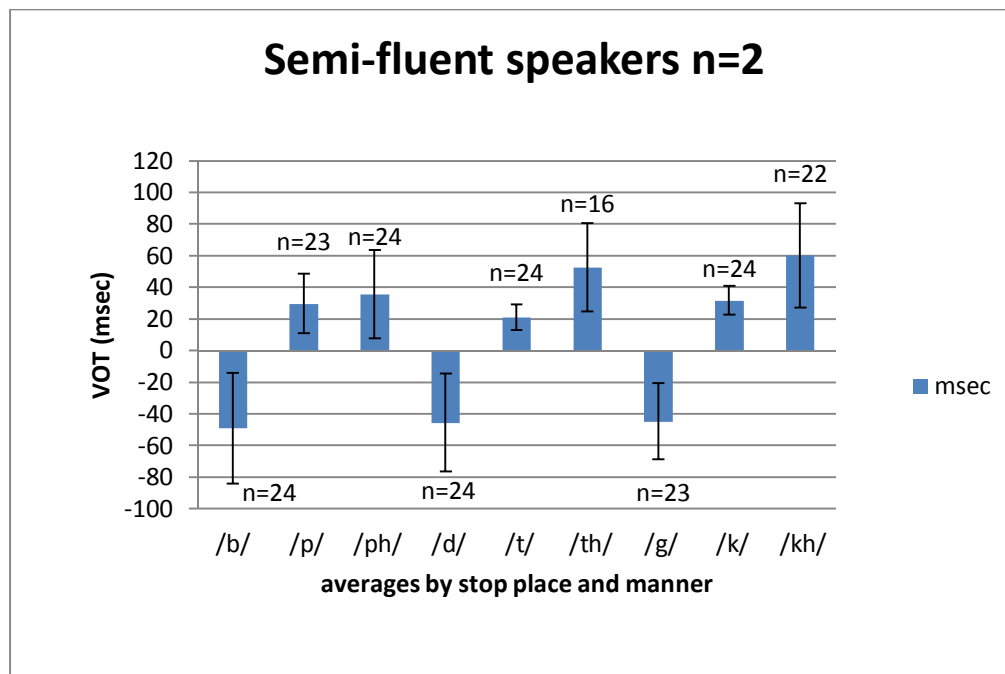


Figure 26.

Talkers E and F (semifluent): mean voice onset time for each target consonant individually; the number of tokens contributing to each mean is presented above/below the bar.

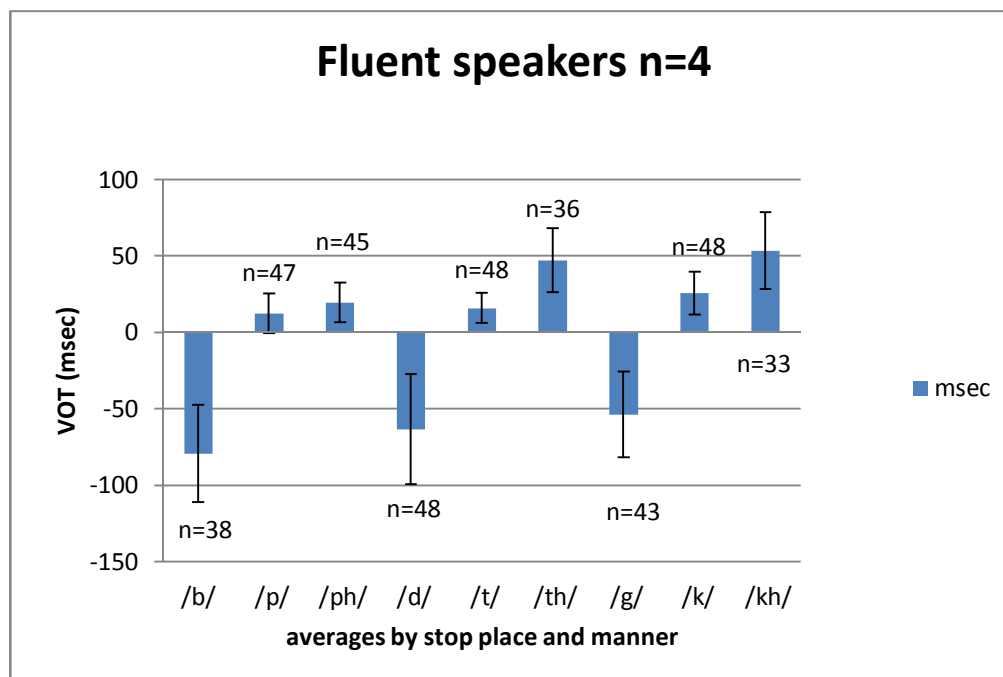


Figure 27.

Talkers A-D (fluent): mean voice onset time for each target consonant individually; the number of tokens contributing to each mean is presented above/below the bar.

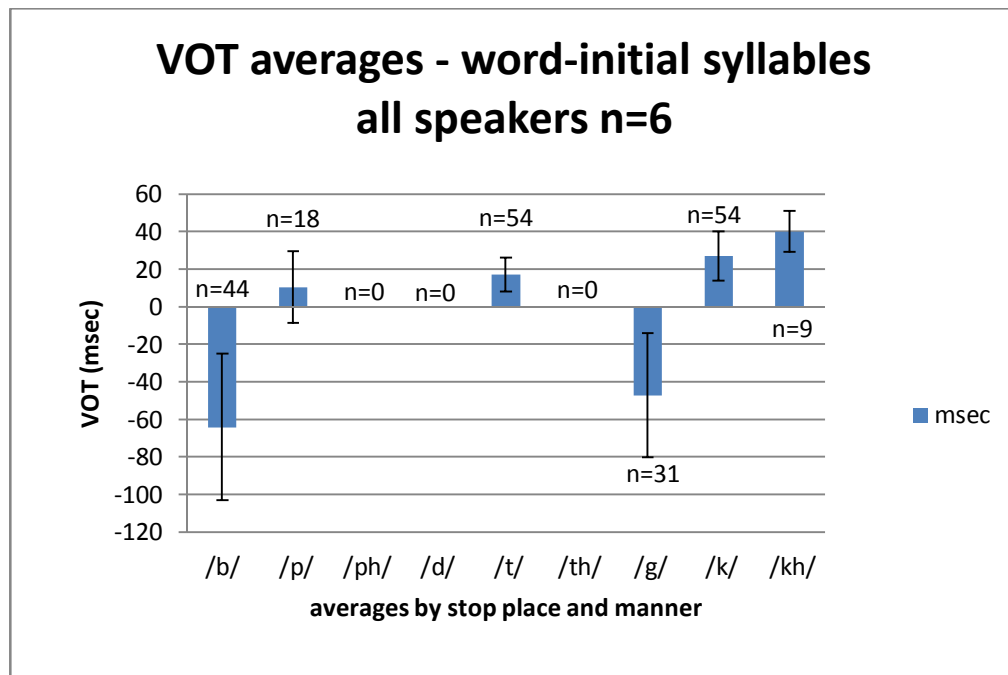


Figure 28.

Talkers A-F (all talkers): mean voice onset time for word initial syllables for each target consonant individually; the number of tokens contributing to each mean is presented above/below the bar. Missing bars indicate that the talker did not produce any usable tokens for that target consonant.



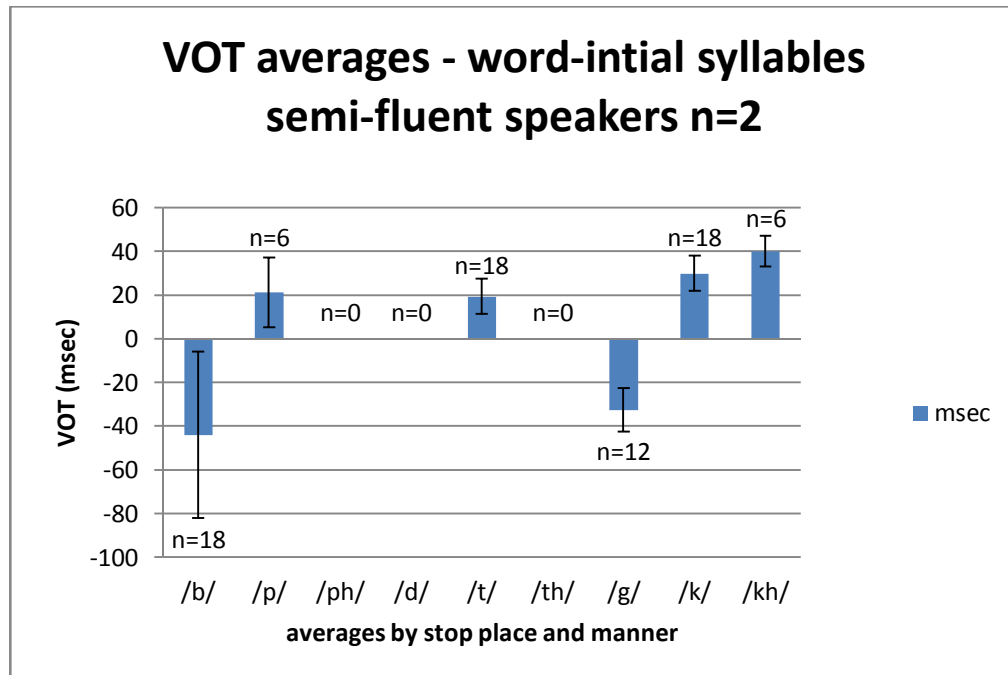


Figure 29.

Talkers E and F (semifluent): mean voice onset time for word initial syllables for each target consonant individually; the number of tokens contributing to each mean is presented above/below the bar. Missing bars indicate that the talker did not produce any usable tokens for that target consonant.

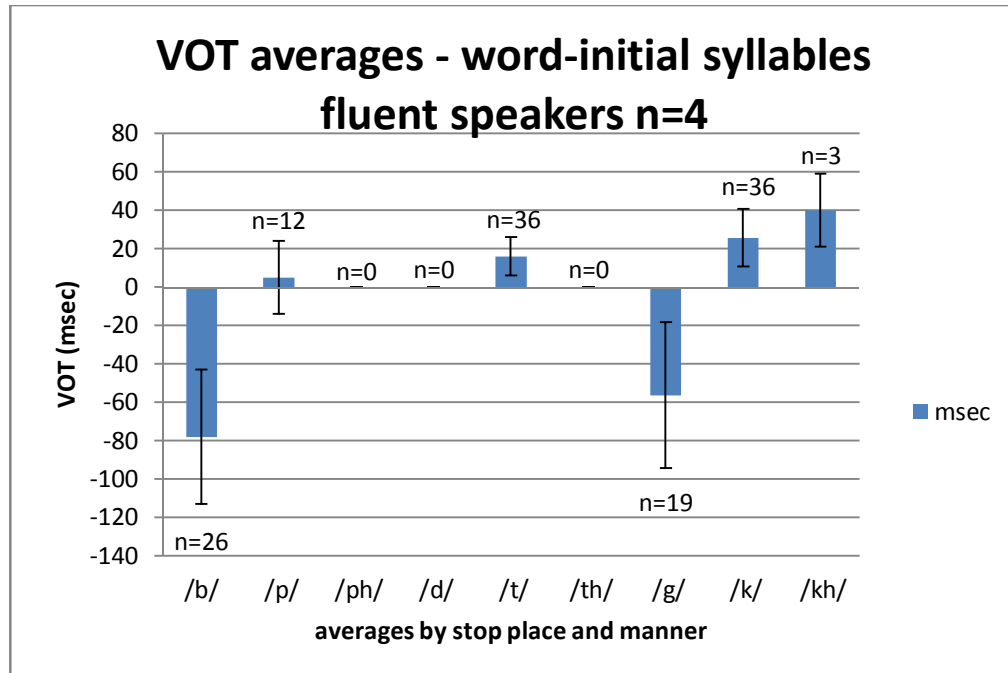


Figure 30.

Talkers A-D (fluent): mean voice onset time for word initial syllables for each target consonant individually; the number of tokens contributing to each mean is presented above/below the bar. Missing bars indicate that the talker did not produce any usable tokens for that target consonant.

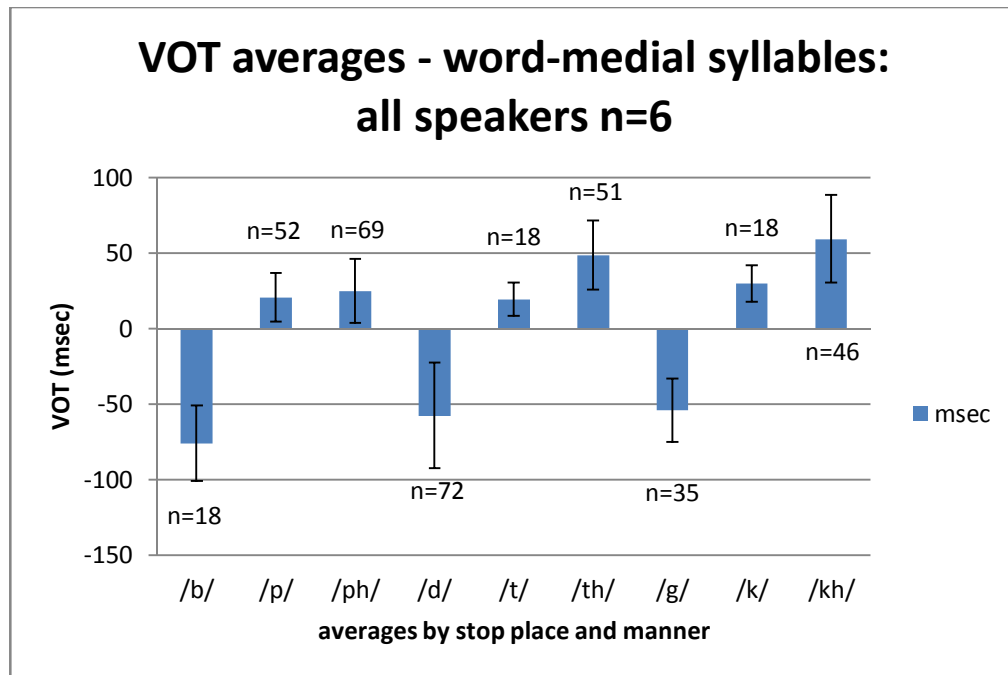


Figure 31.

Talkers A-F (all talkers): mean voice onset time for word medial syllables for each target consonant individually; the number of tokens contributing to each mean is presented above/below the bar.

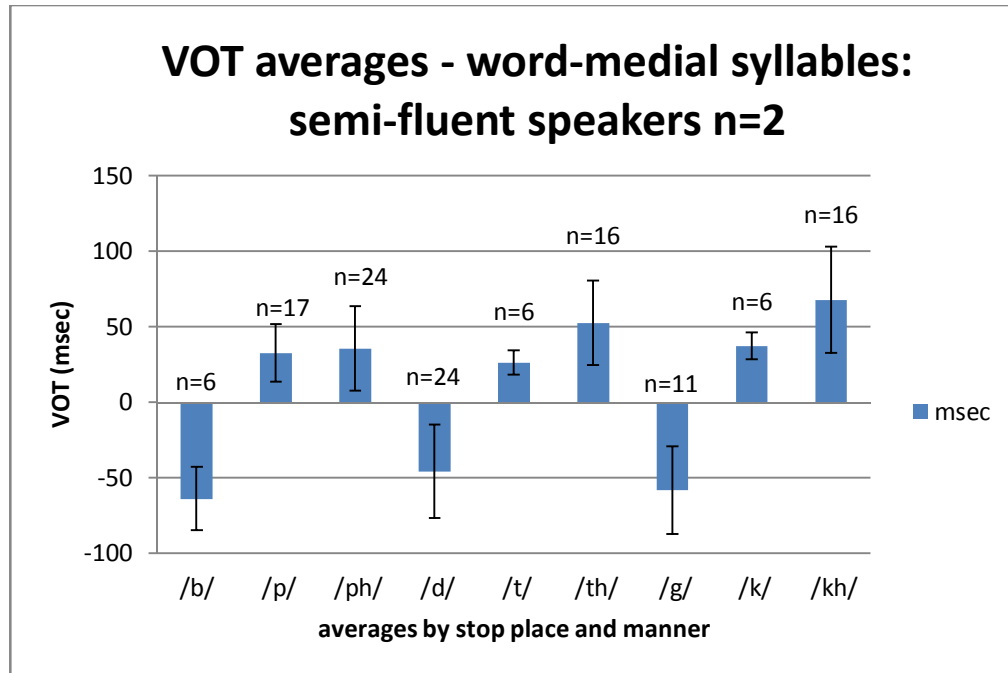


Figure 32.

Talkers E and F (semifluent): mean voice onset time for word medial syllables for each target consonant individually; the number of tokens contributing to each mean is presented above/below the bar.

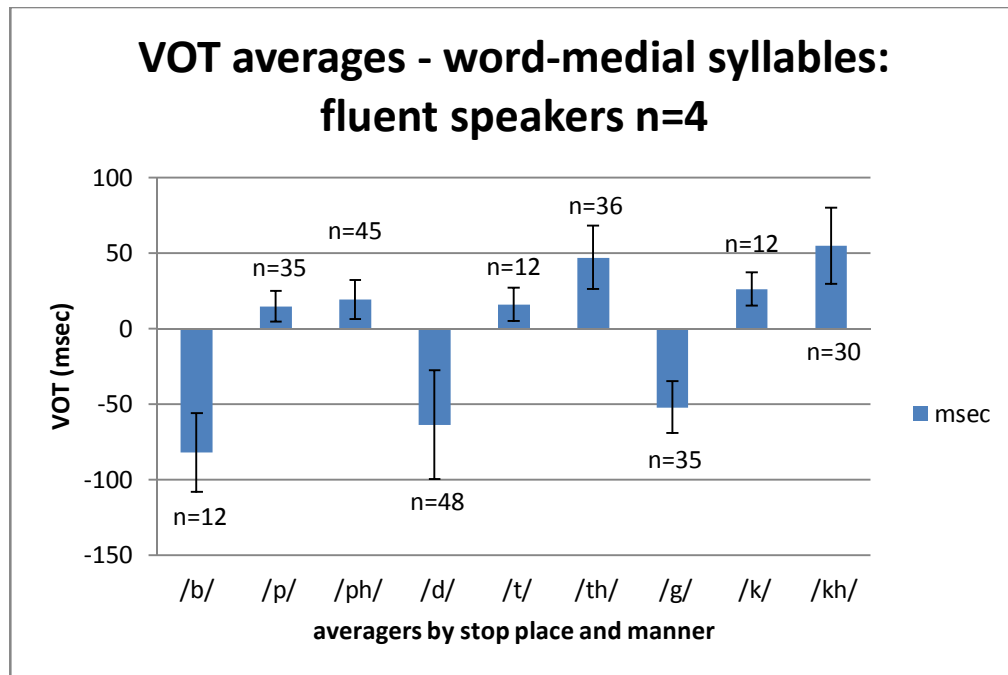


Figure 33.

Talkers A-D (fluent): mean voice onset time for word medial syllables for each target consonant individually; the number of tokens contributing to each mean is presented above/below the bar.

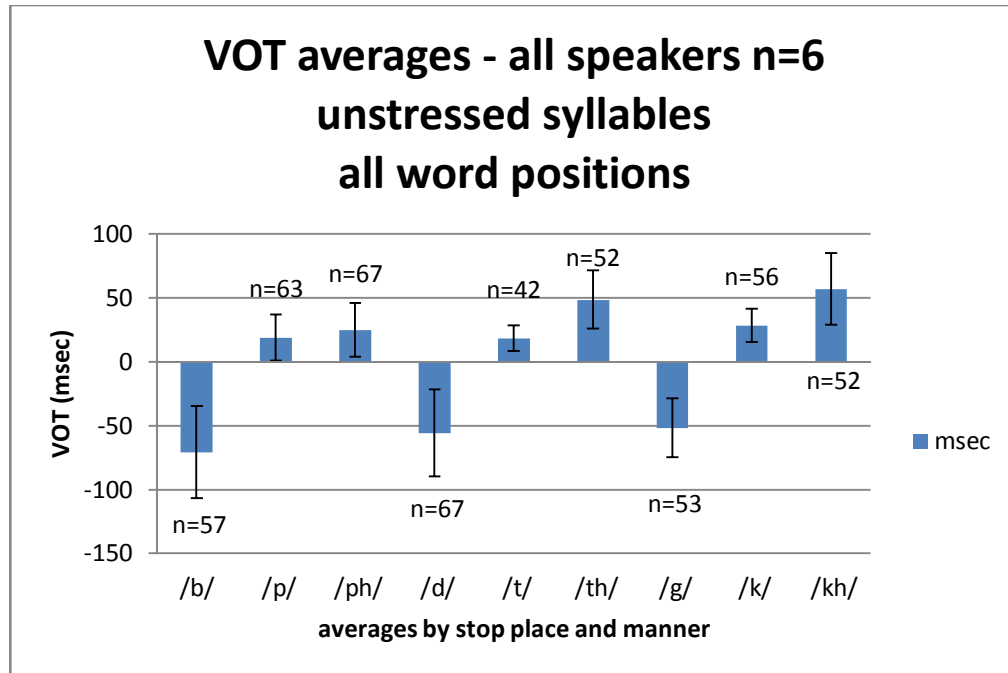


Figure 34.

Talkers A-F (all talkers): mean voice onset time for unstressed syllables in all word positions for each target consonant individually; the number of tokens contributing to each mean is presented above/below the bar.

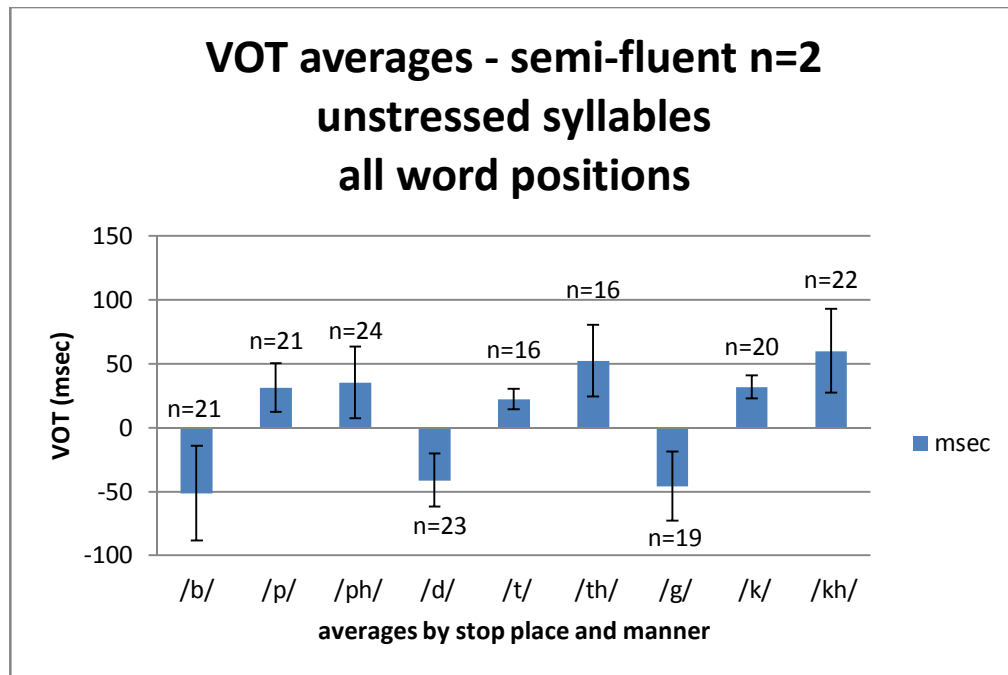


Figure 35.

Talkers E and F (semifluent): mean voice onset time for unstressed syllables in all word positions for each target consonant individually; the number of tokens contributing to each mean is presented above/below the bar.

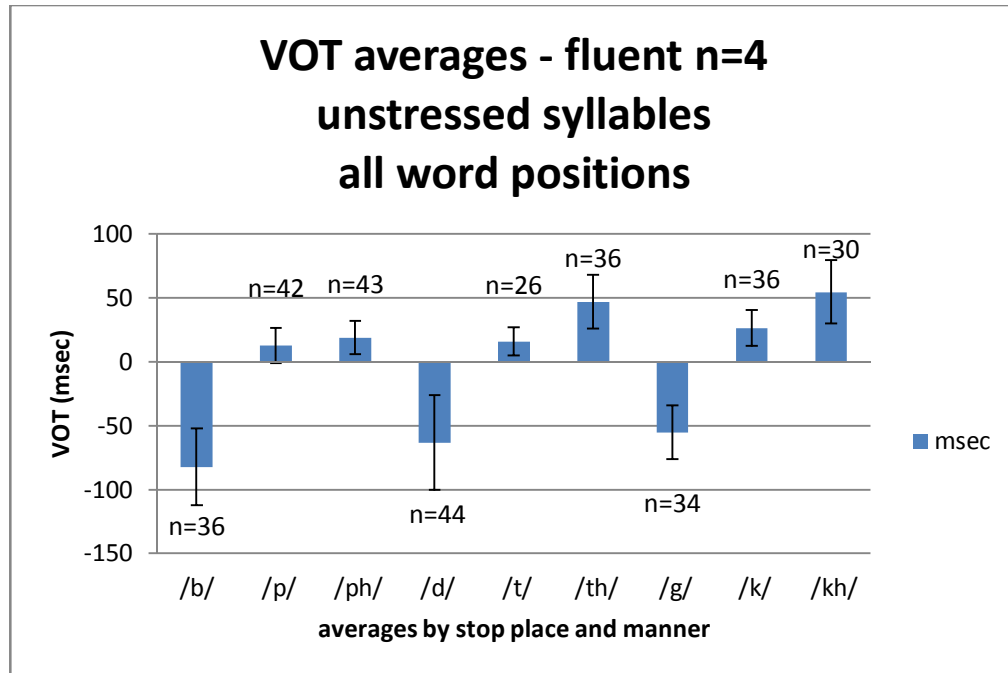


Figure 36.

Talkers A-D (fluent): mean voice onset time for unstressed syllables in all word positions for each target consonant individually; the number of tokens contributing to each mean is presented above/below the bar.



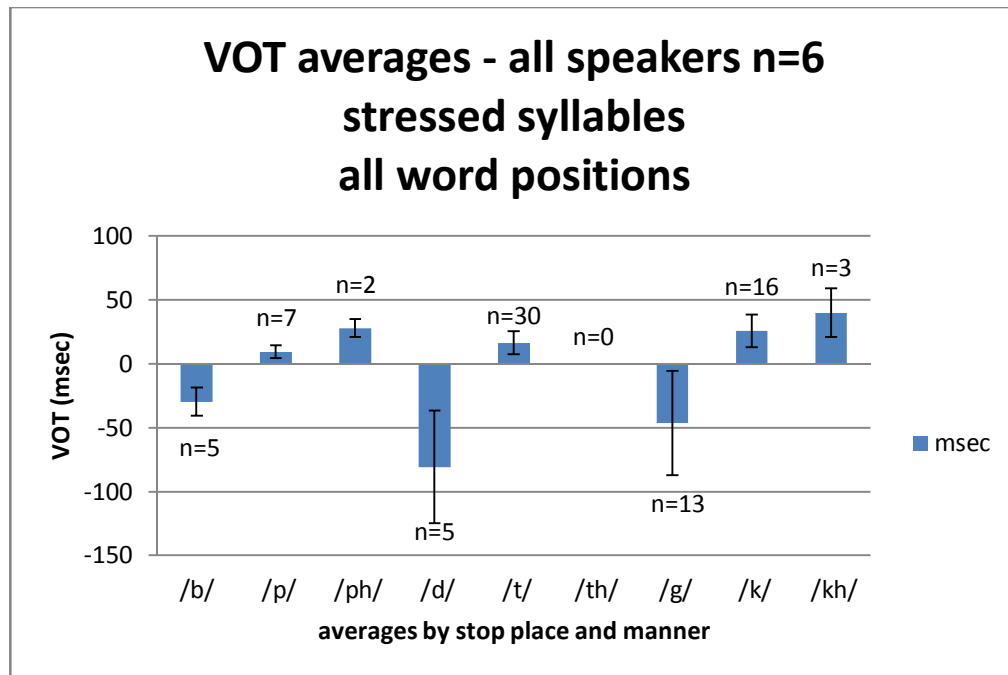


Figure 37.

Talkers A-F (all talkers): mean voice onset time for stressed syllables in all word positions for each target consonant individually; the number of tokens contributing to each mean is presented above/below the bar. Missing bars indicate that the talker did not produce any usable tokens for that target consonant.

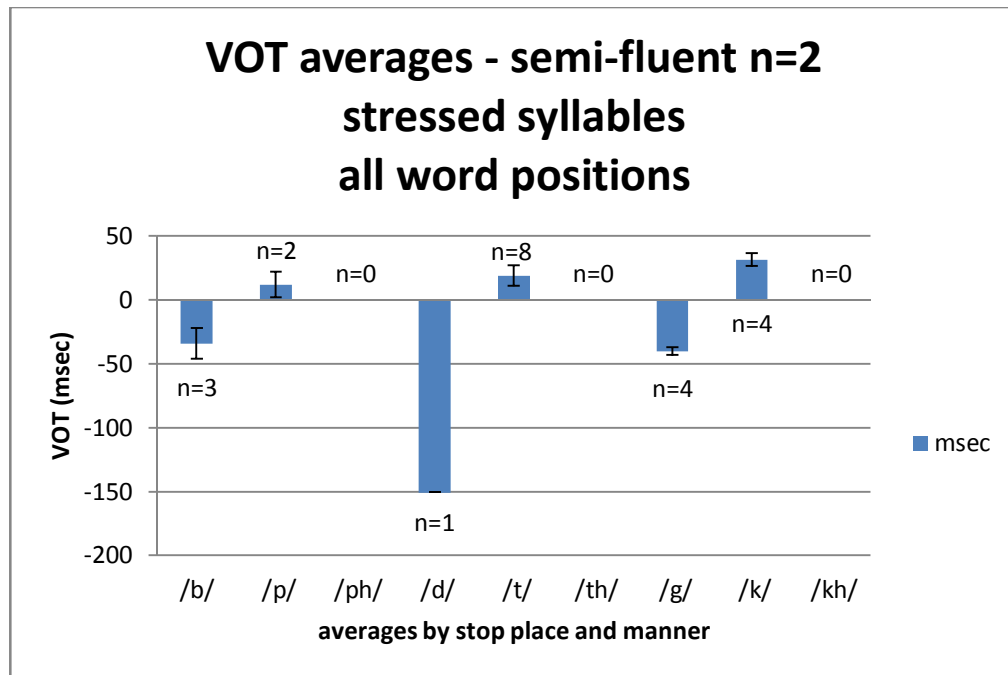


Figure 38.

Talkers E and F (semifluent): mean voice onset time for stressed syllables in all word positions for each target consonant individually; the number of tokens contributing to each mean is presented above/below the bar. Missing bars indicate that the talker did not produce any usable tokens for that target consonant.

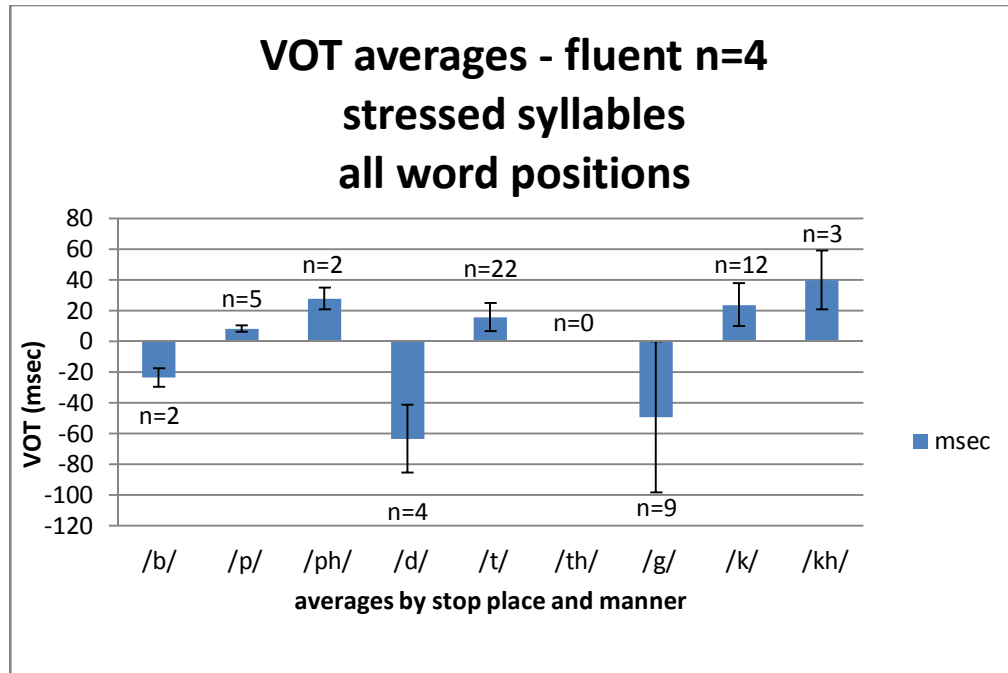


Figure 39.

Talkers A-D (fluent): mean voice onset time for stressed syllables in all word positions for each target consonant individually; the number of tokens contributing to each mean is presented above/below the bar. Missing bars indicate that the talker did not produce any usable tokens for that target consonant.

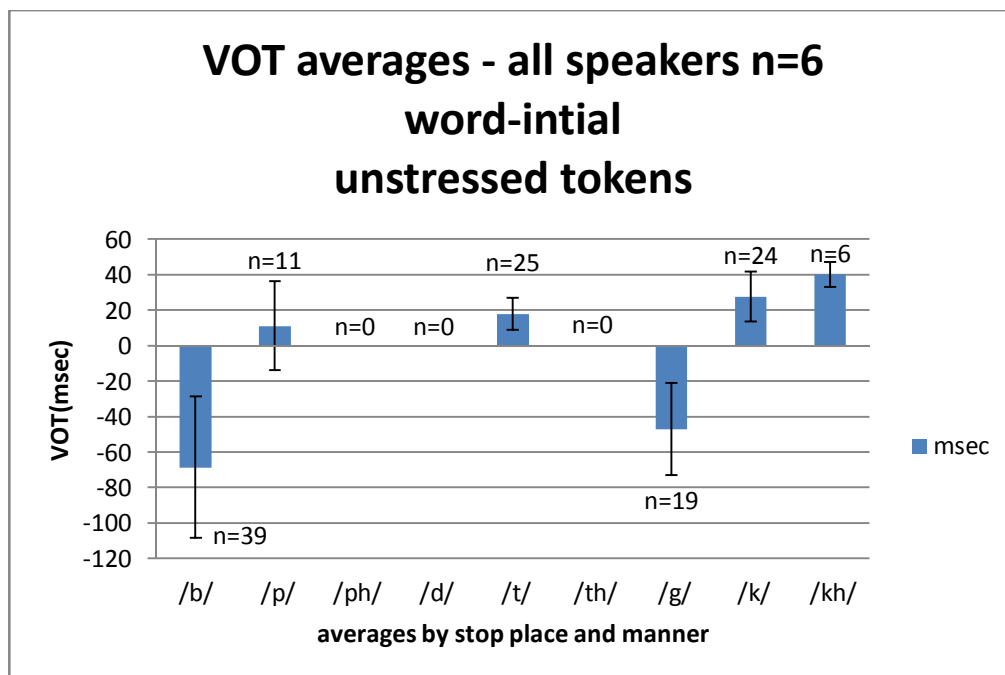


Figure 40.

Talkers A-F (all talkers): mean voice onset time for word initial unstressed syllables for each target consonant individually; the number of tokens contributing to each mean is presented above/below the bar. Missing bars indicate that the talker did not produce any usable tokens for that target consonant.

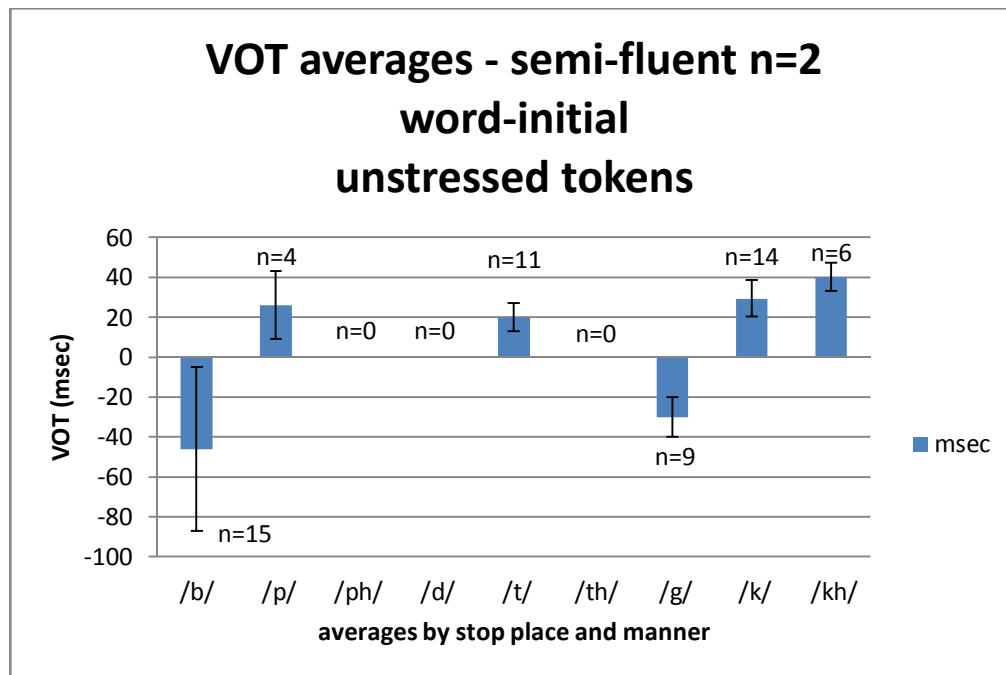


Figure 41.

Talkers E and F (semifluent): mean voice onset time for word initial unstressed syllables for each target consonant individually; the number of tokens contributing to each mean is presented above/below the bar. Missing bars indicate that the talker did not produce any usable tokens for that target consonant.

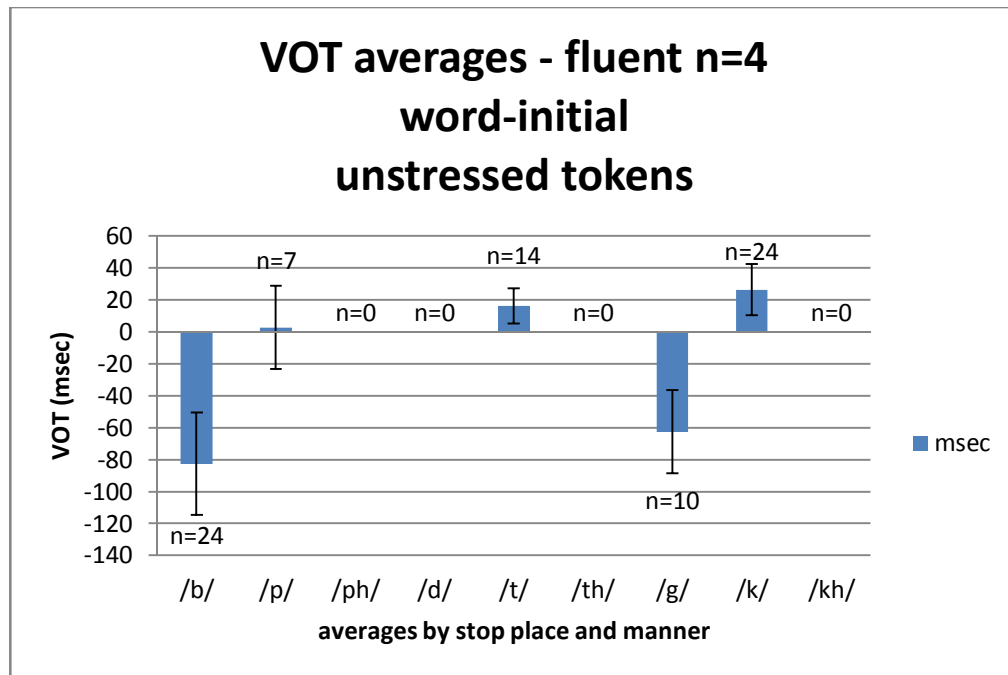


Figure 42.

Talkers A-D (fluent): mean voice onset time for word initial unstressed syllables for each target consonant individually; the number of tokens contributing to each mean is presented above/below the bar. Missing bars indicate that the talker did not produce any usable tokens for that target consonant.

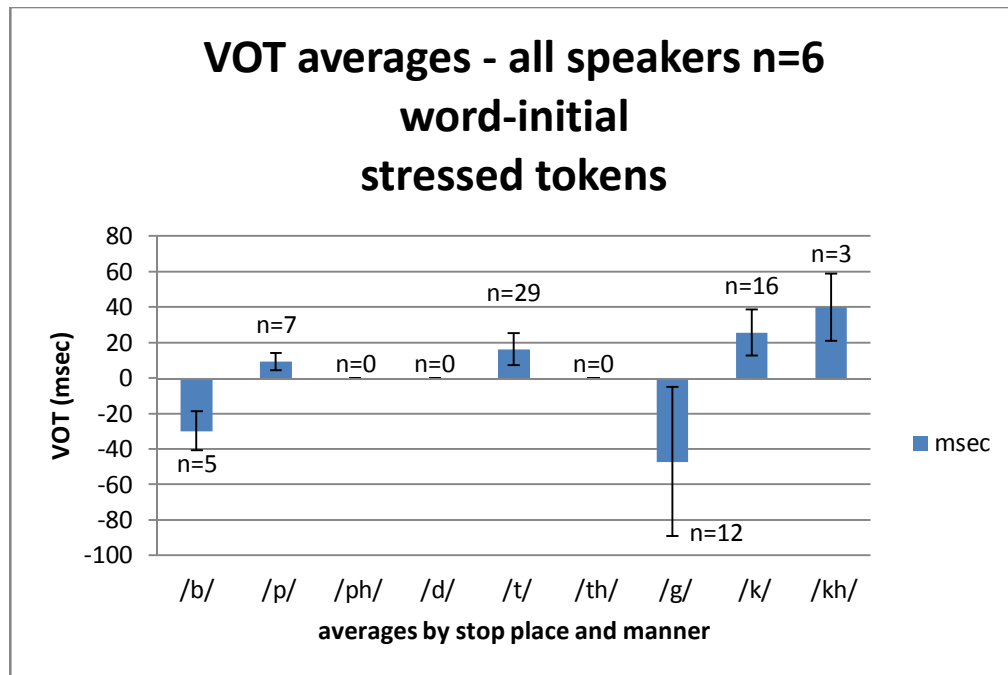


Figure 43.

Talkers A-F (all talkers): mean voice onset time for word initial stressed syllables for each target consonant individually; the number of tokens contributing to each mean is presented above/below the bar. Missing bars indicate that the talker did not produce any usable tokens for that target consonant.

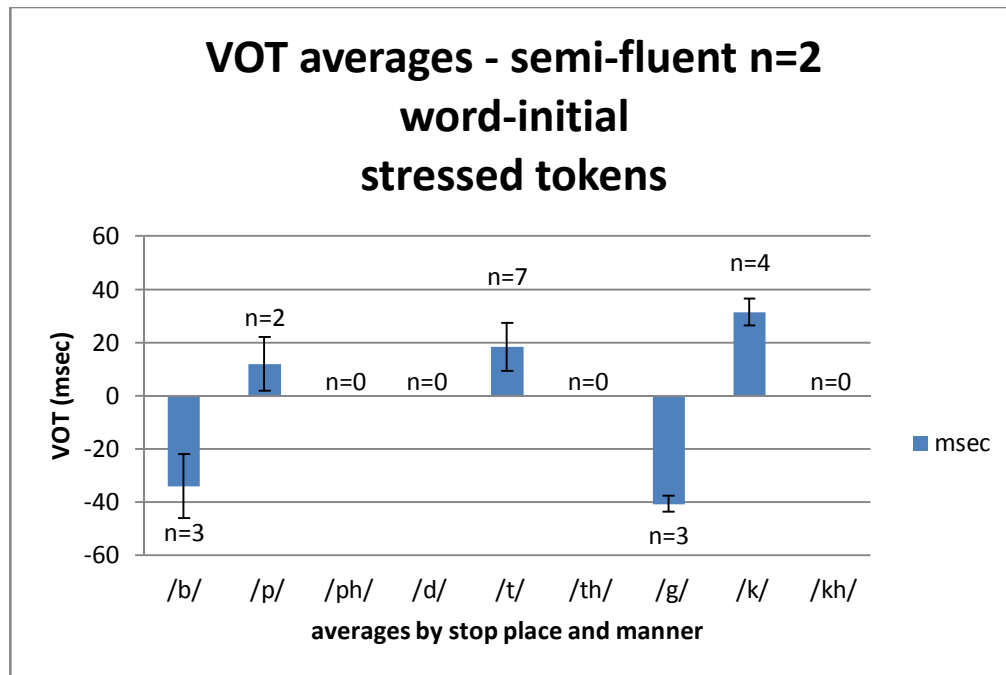


Figure 44.

Talkers E and F (semifluent): mean voice onset time for word initial stressed syllables for each target consonant individually; the number of tokens contributing to each mean is presented above/below the bar. Missing bars indicate that the talker did not produce any usable tokens for that target consonant.



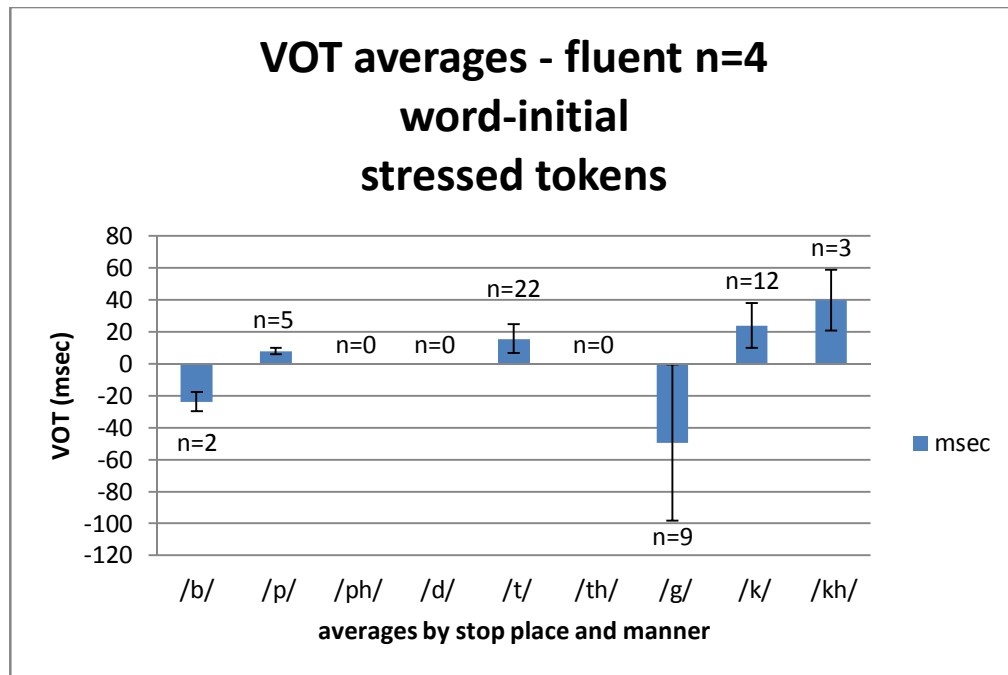


Figure 45.

Talkers A-D (fluent): mean voice onset time for word initial stressed syllables for each target consonant individually; the number of tokens contributing to each mean is presented above/below the bar. Missing bars indicate that the talker did not produce any usable tokens for that target consonant.

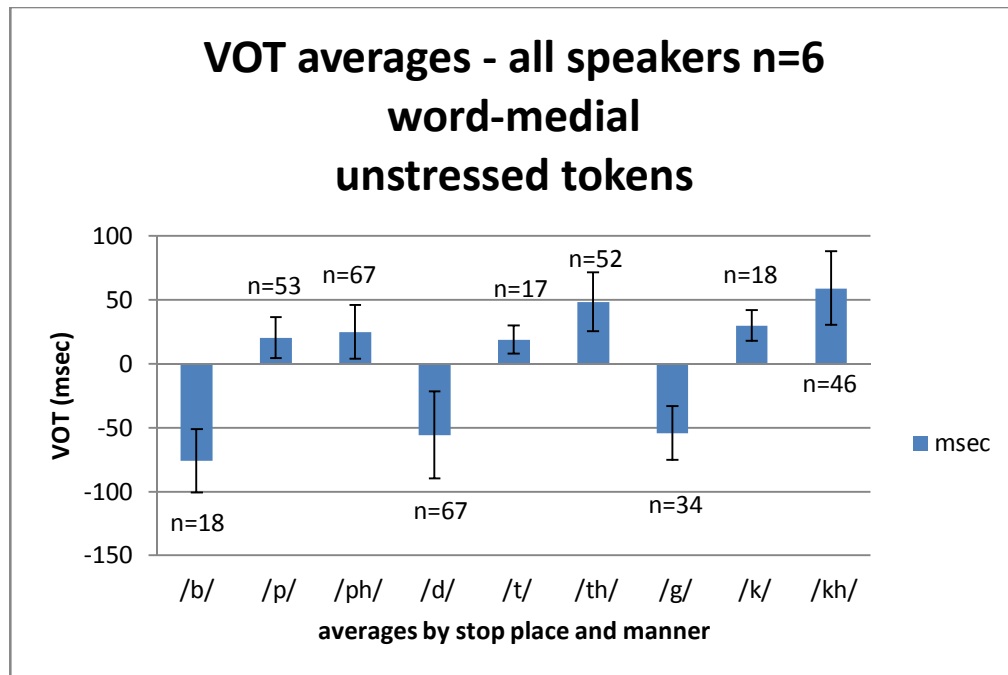


Figure 46.

Talkers A-F (all talkers): mean voice onset time for word medial unstressed syllables for each target consonant individually; the number of tokens contributing to each mean is presented above/below the bar.

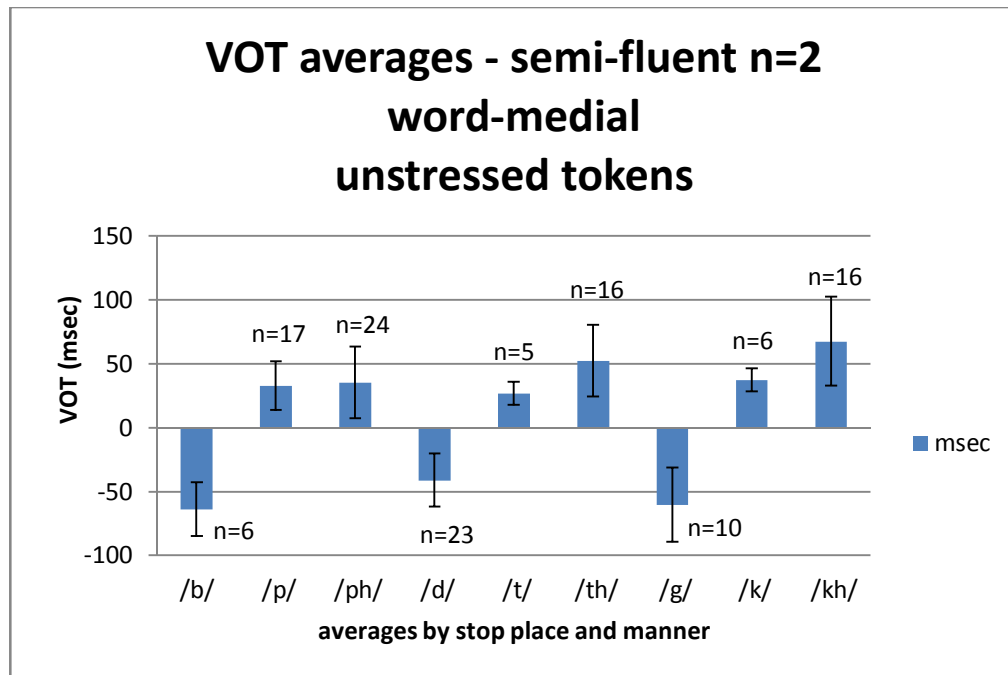


Figure 47.

Talkers E and F (semifluent): mean voice onset time for word medial unstressed syllables for each target consonant individually; the number of tokens contributing to each mean is presented above/below the bar.

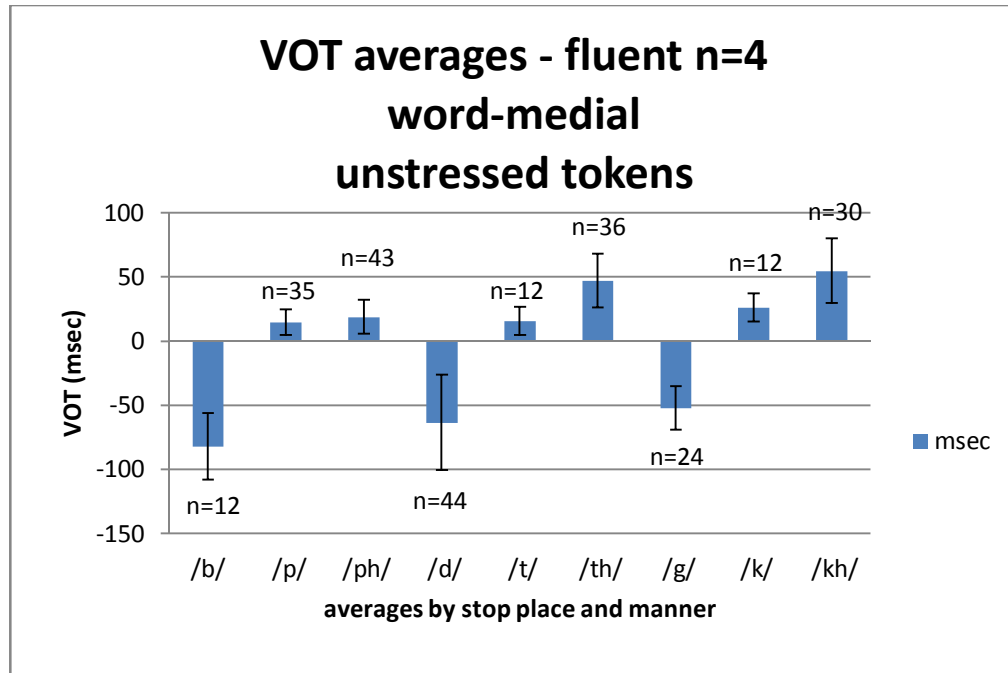


Figure 48.

Talkers A-D (fluent): mean voice onset time for word medial unstressed syllables for each target consonant individually; the number of tokens contributing to each mean is presented above/below the bar.

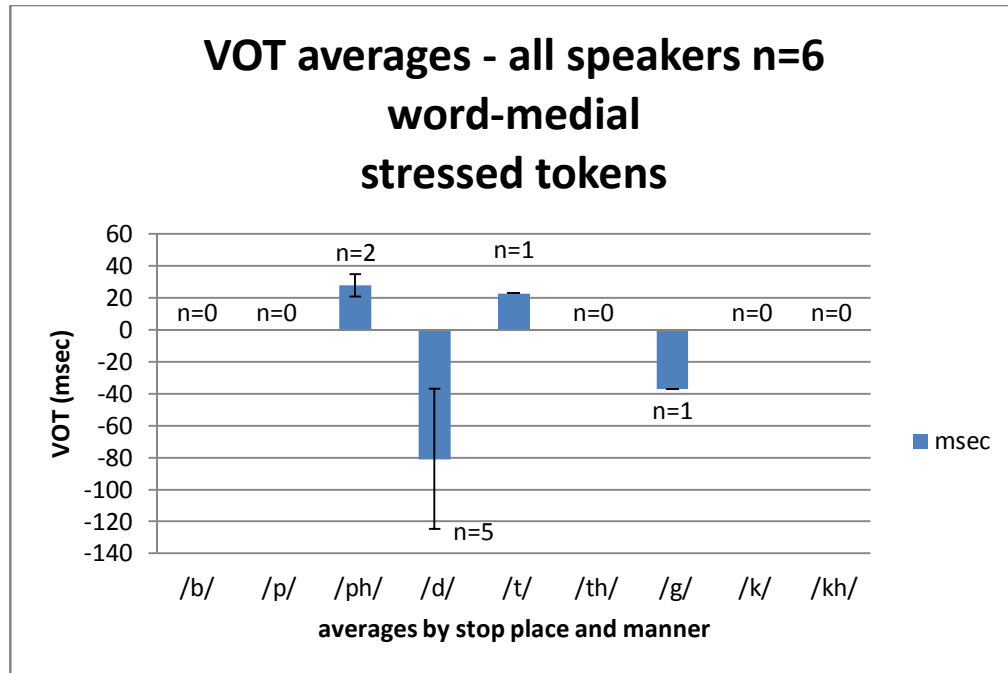


Figure 49.

Talkers A-F (all talkers): mean voice onset time for word medial stressed syllables for each target consonant individually; the number of tokens contributing to each mean is presented above/below the bar. Missing bars indicate that the talker did not produce any usable tokens for that target consonant.

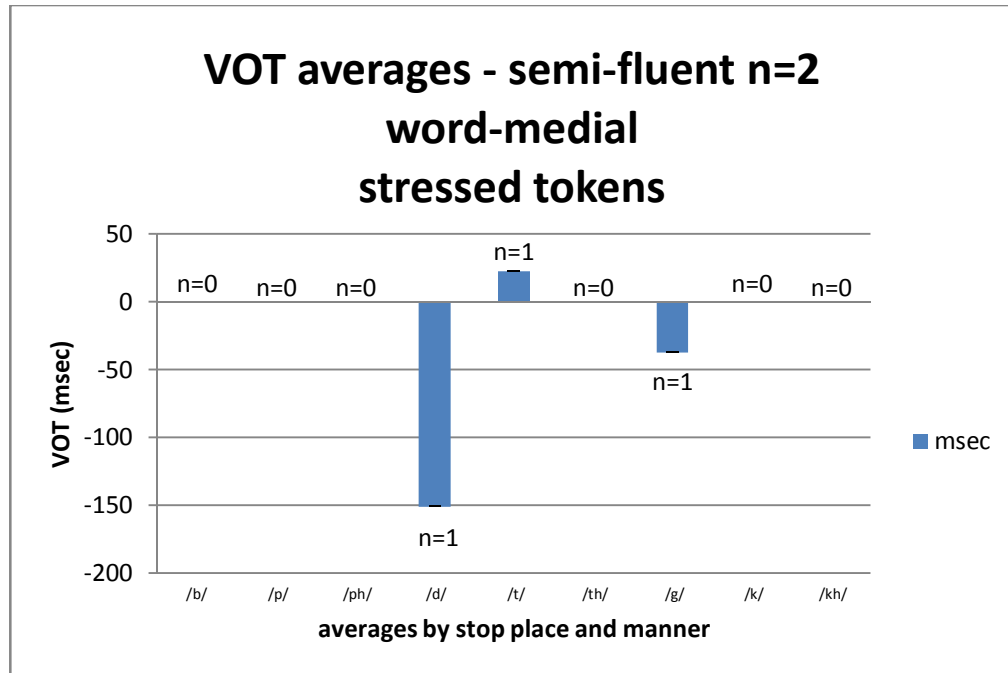


Figure 50.

Talkers E and F (semifluent): mean voice onset time for word medial stressed syllables for each target consonant individually; the number of tokens contributing to each mean is presented above/below the bar. Missing bars indicate that the talker did not produce any usable tokens for that target consonant.

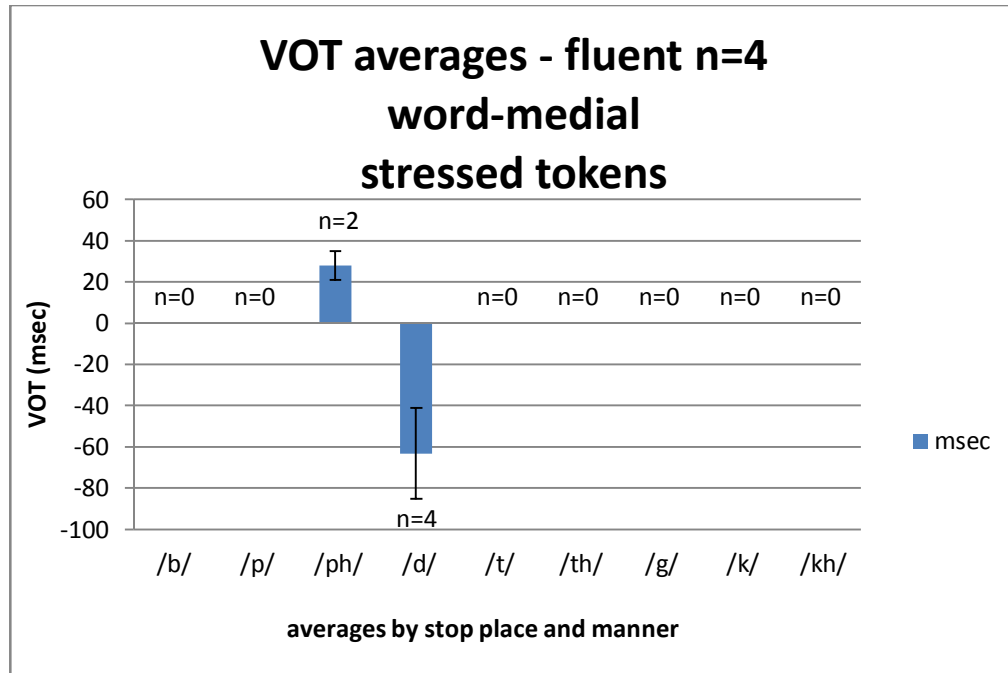


Figure 51.

Talkers A-D (fluent): mean voice onset time for word medial stressed syllables for each target consonant individually; the number of tokens contributing to each mean is presented above/below the bar. Missing bars indicate that the talker did not produce any usable tokens for that target consonant.

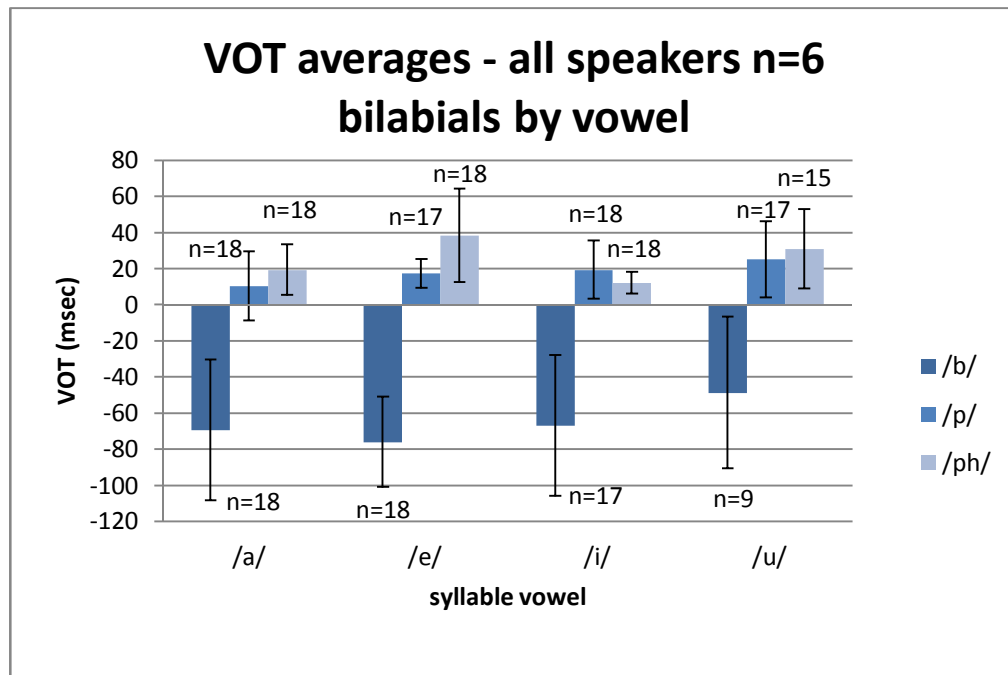


Figure 52.

Talkers A-F (all talkers): mean voice onset time for bilabials by individual vowel; the number of tokens contributing to each mean is presented above/below the bar.



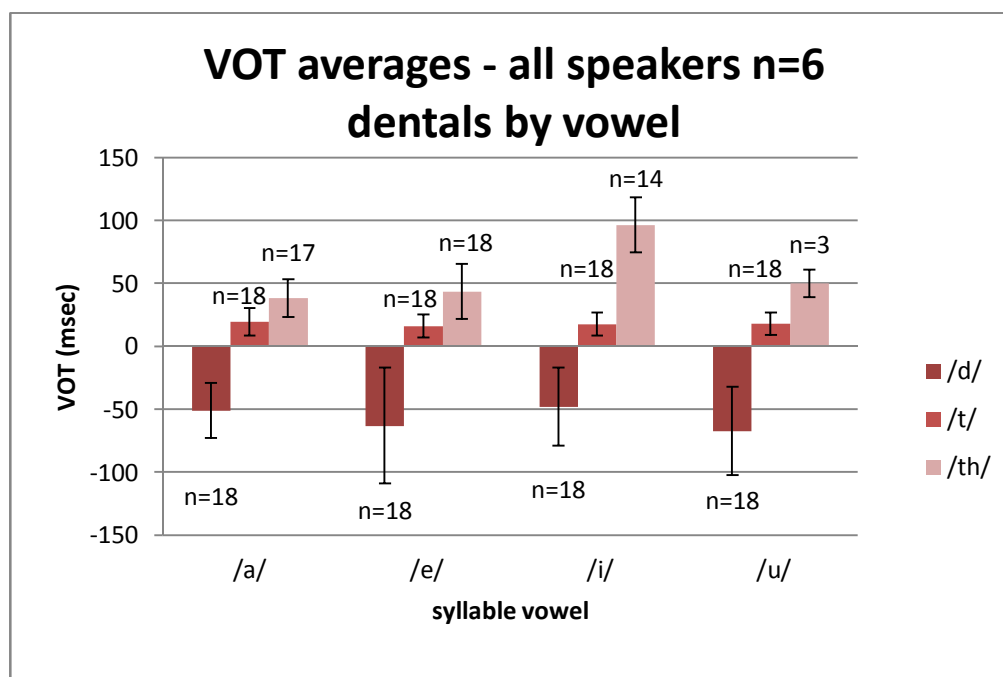


Figure 53.

Talkers A-F (all talkers): mean voice onset time for dentals by individual vowel; the number of tokens contributing to each mean is presented above/below the bar.

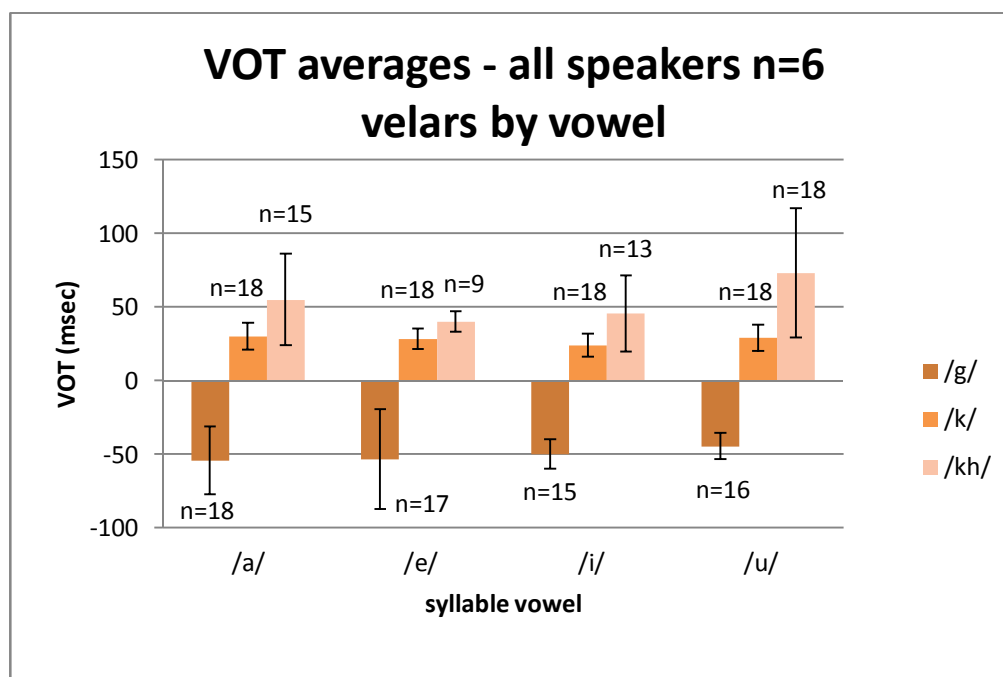


Figure 54.

Talkers A-F (all talkers): mean voice onset time for velars by individual vowel; the number of tokens contributing to each mean is presented above/below the bar.

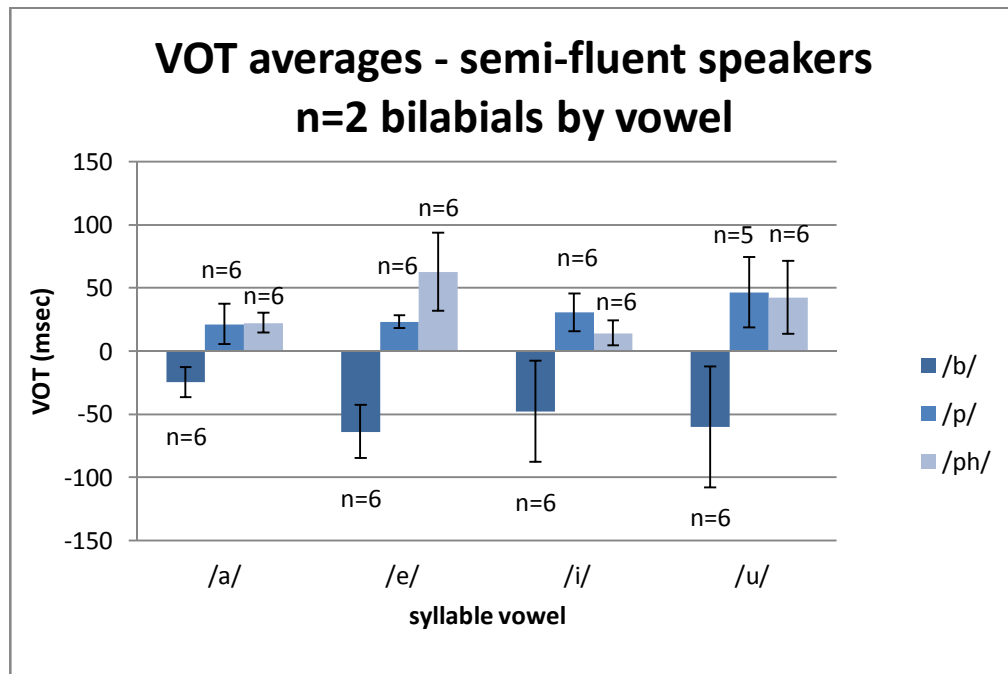


Figure 55.

Talkers E and F (semifluent): mean voice onset time for bilabials by individual vowel; the number of tokens contributing to each mean is presented above/below the bar.

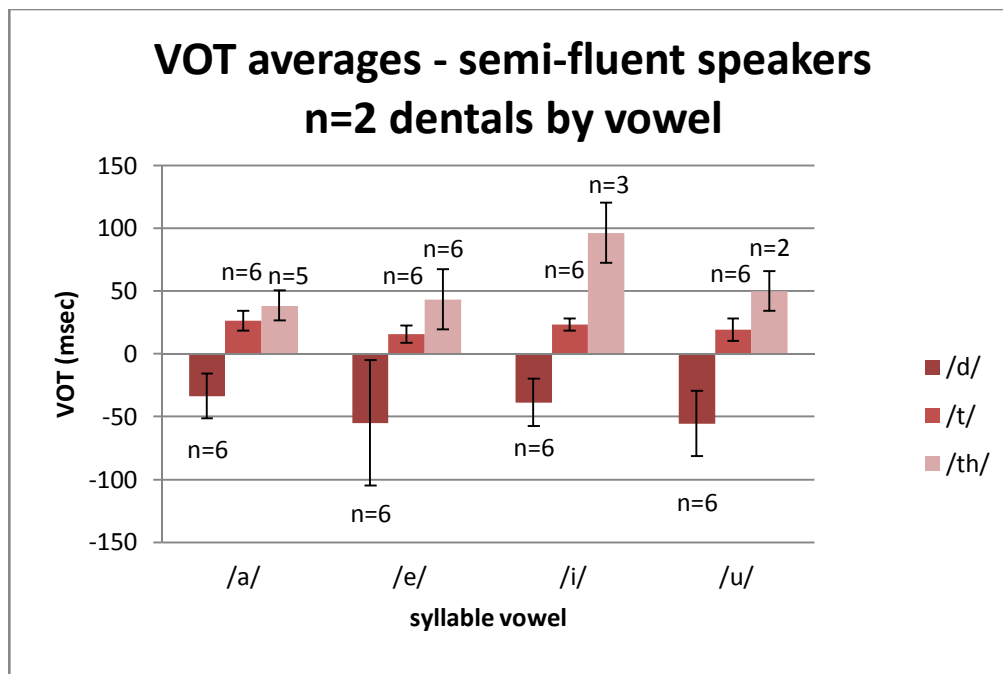


Figure 56.

Talkers E and F (semifluent): mean voice onset time for dentals by individual vowel; the number of tokens contributing to each mean is presented above/below the bar.

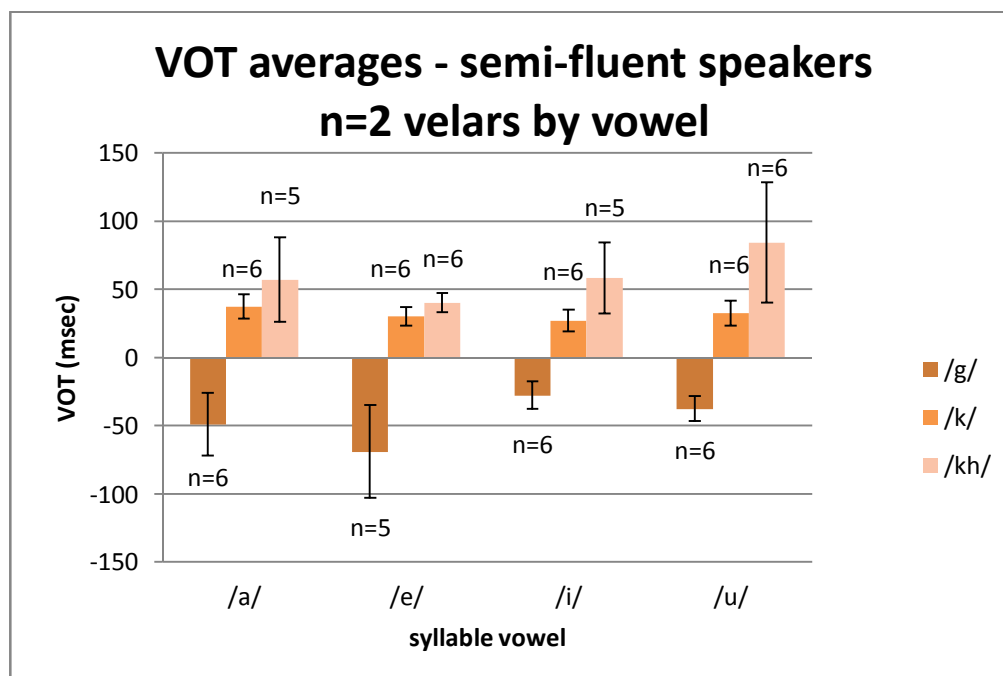


Figure 57.

Talkers E and F (semifluent): mean voice onset time for velars by individual vowel; the number of tokens contributing to each mean is presented above/below the bar.

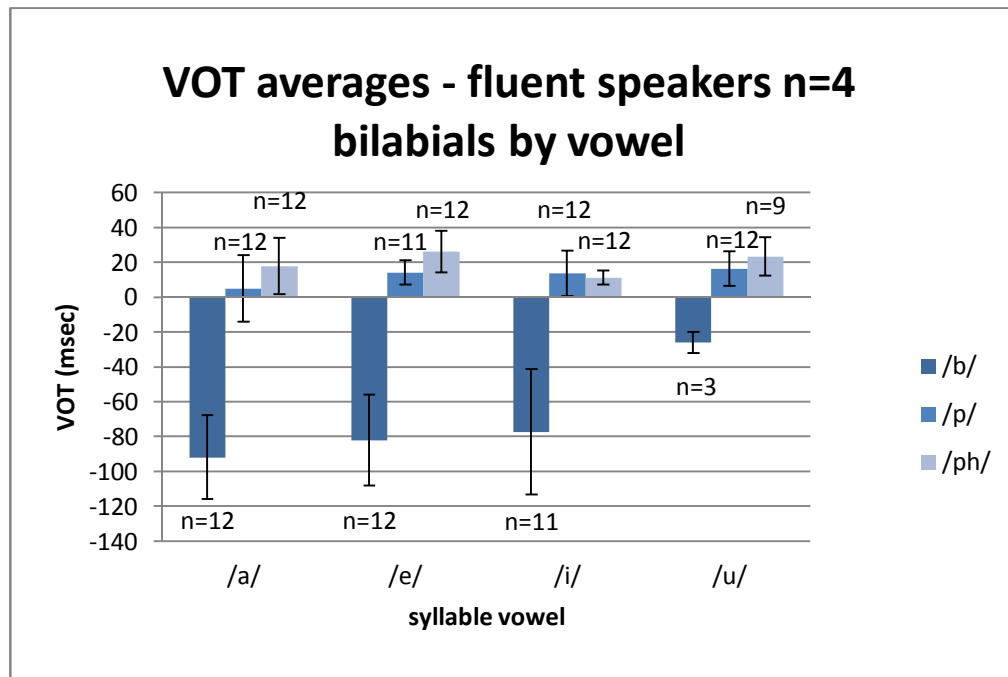


Figure 58.

Talkers A-D (fluent): mean voice onset time for bilabials by individual vowel; the number of tokens contributing to each mean is presented above/below the bar.

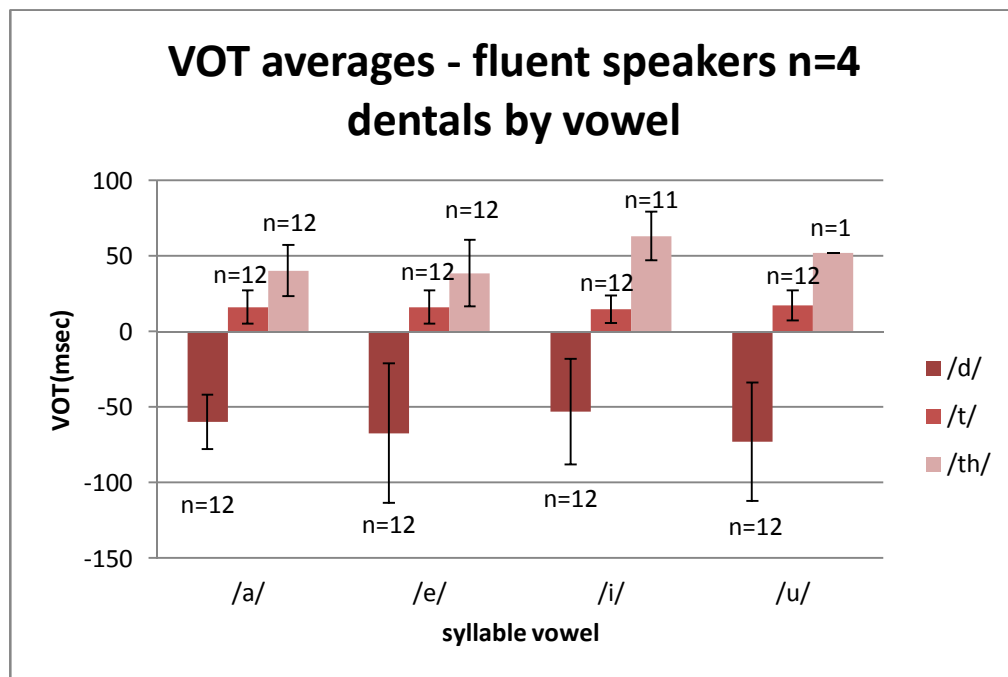


Figure 59.

Talkers A-D (fluent): mean voice onset time for dentals by individual vowel; the number of tokens contributing to each mean is presented above/below the bar.

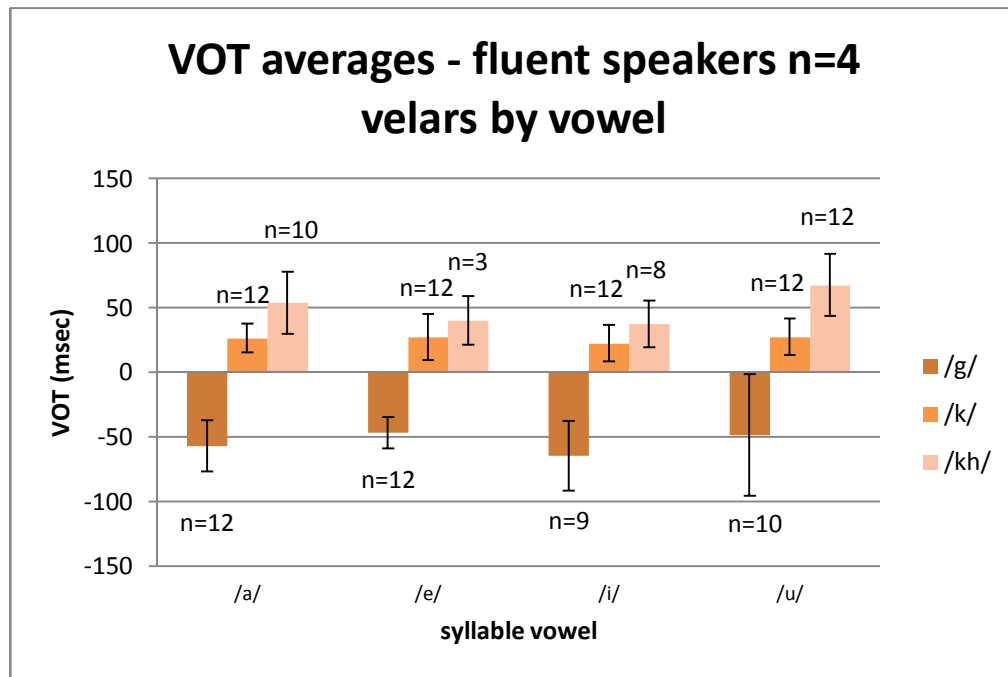


Figure 60.

Talkers A-D (fluent): mean voice onset time for velars by individual vowel; the number of tokens contributing to each mean is presented above/below the bar.



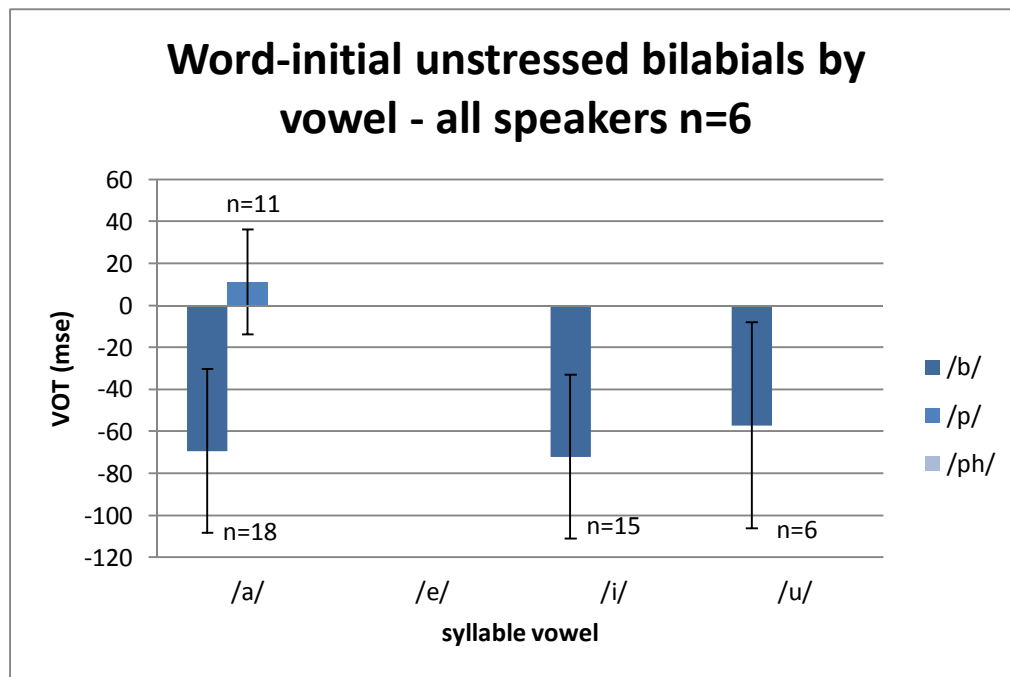


Figure 61.

Talkers A-F (all talkers): mean voice onset time for word initial unstressed bilabials by individual vowel; the number of tokens contributing to each mean is presented above/below the bar. Missing bars indicate that the talker did not produce any usable tokens for that target consonant.

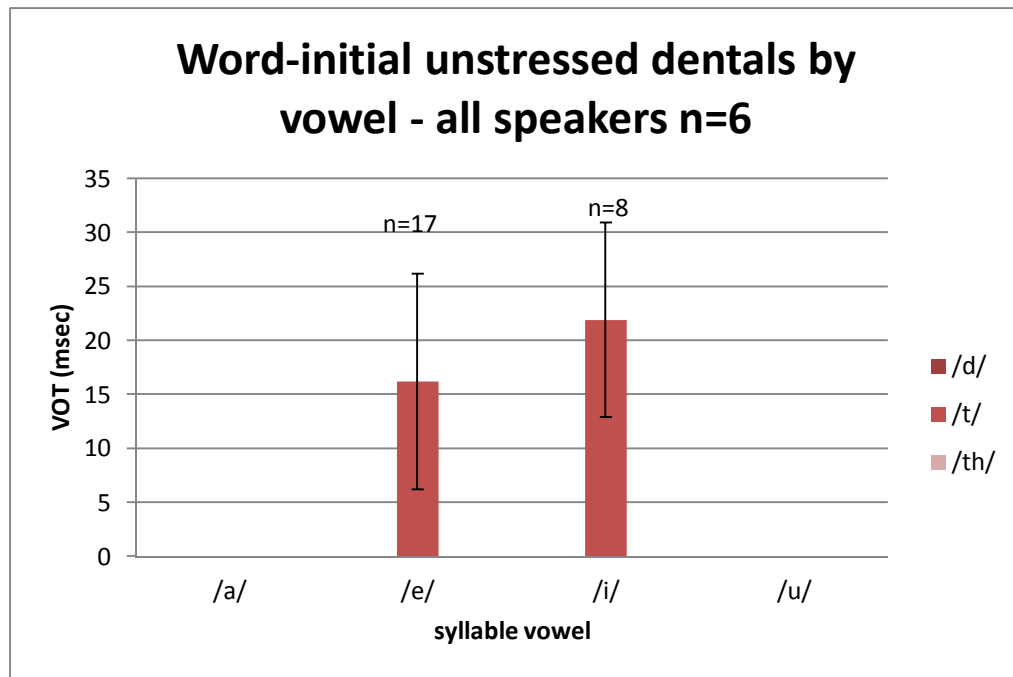


Figure 62.

Talkers A-F (all talkers): mean voice onset time for word initial unstressed dentals by individual vowel; the number of tokens contributing to each mean is presented above/below the bar. Missing bars indicate that the talker did not produce any usable tokens for that target consonant.

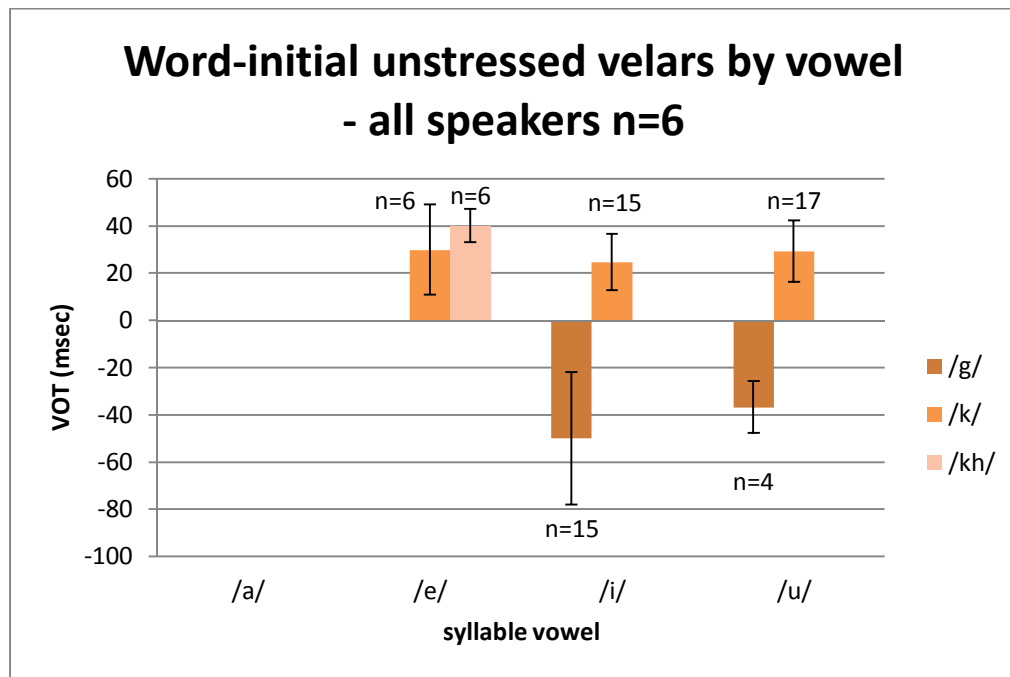


Figure 63.

Talkers A-F (all talkers): mean voice onset time for word initial unstressed velars by individual vowel; the number of tokens contributing to each mean is presented above/below the bar. Missing bars indicate that the talker did not produce any usable tokens for that target consonant.

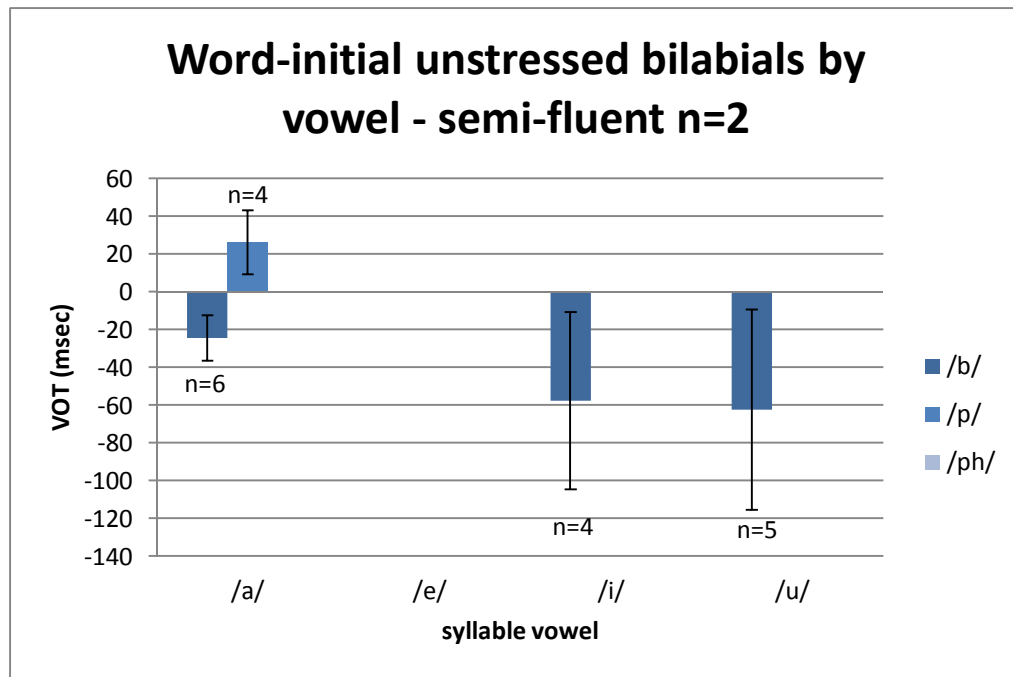


Figure 64.

Talkers E and F (semifluent): mean voice onset time for word initial unstressed bilabials by individual vowel; the number of tokens contributing to each mean is presented above/below the bar. Missing bars indicate that the talker did not produce any usable tokens for that target consonant.

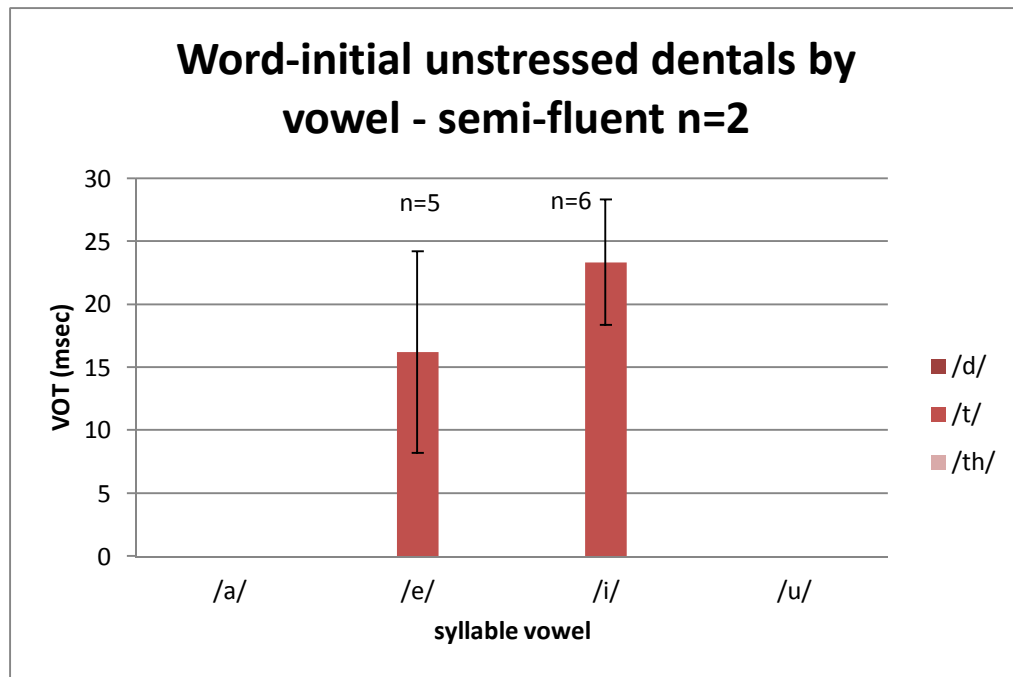


Figure 65.

Talkers E and F (semifluent): mean voice onset time for word initial unstressed dentals by individual vowel; the number of tokens contributing to each mean is presented above/below the bar. Missing bars indicate that the talker did not produce any usable tokens for that target consonant.

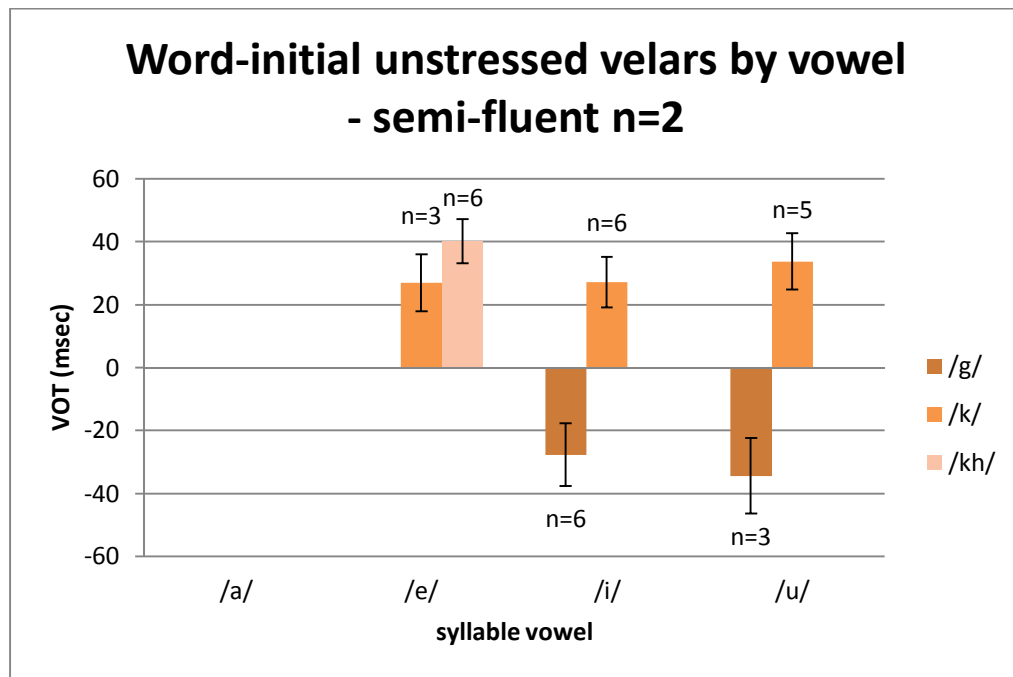


Figure 66.

Talkers E and F (semifluent): mean voice onset time for word initial unstressed velars by individual vowel; the number of tokens contributing to each mean is presented above/below the bar. Missing bars indicate that the talker did not produce any usable tokens for that target consonant.

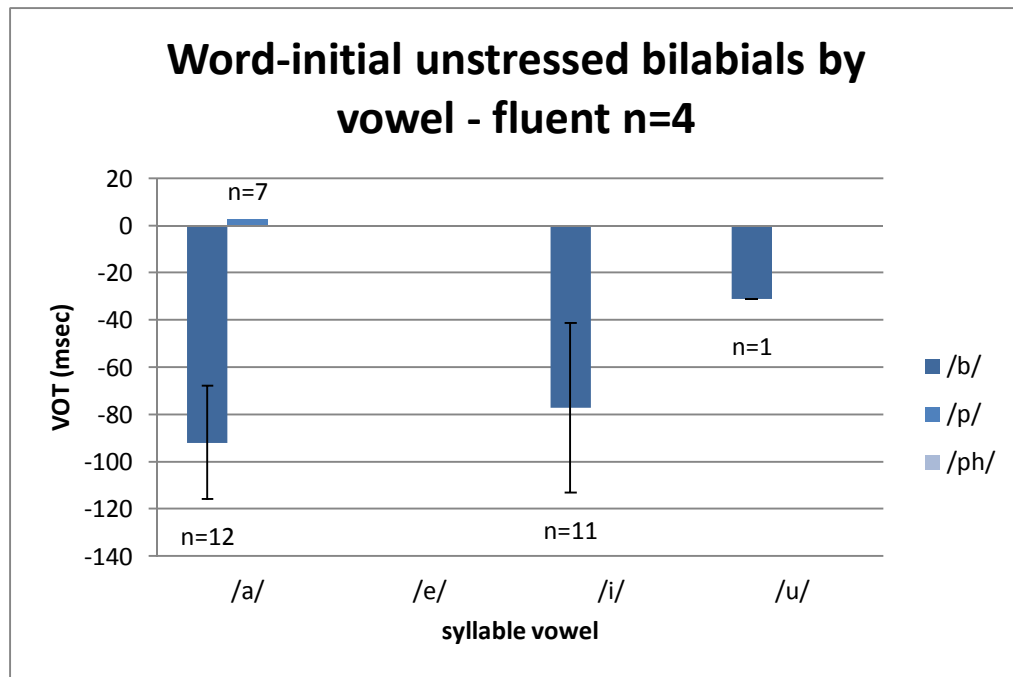


Figure 67.

Talkers A-D (fluent): mean voice onset time for word initial unstressed bilabials by individual vowel; the number of tokens contributing to each mean is presented above/below the bar. Missing bars indicate that the talker did not produce any usable tokens for that target consonant.

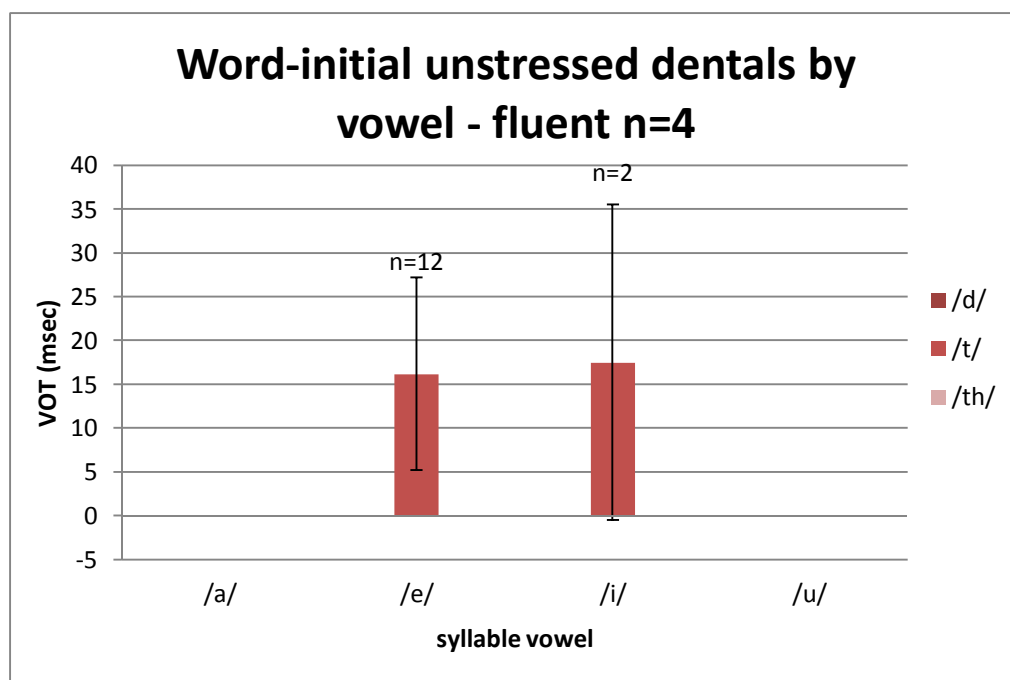


Figure 68.

Talkers A-D (fluent): mean voice onset time for word initial unstressed dentals by individual vowel; the number of tokens contributing to each mean is presented above/below the bar. Missing bars indicate that the talker did not produce any usable tokens for that target consonant.



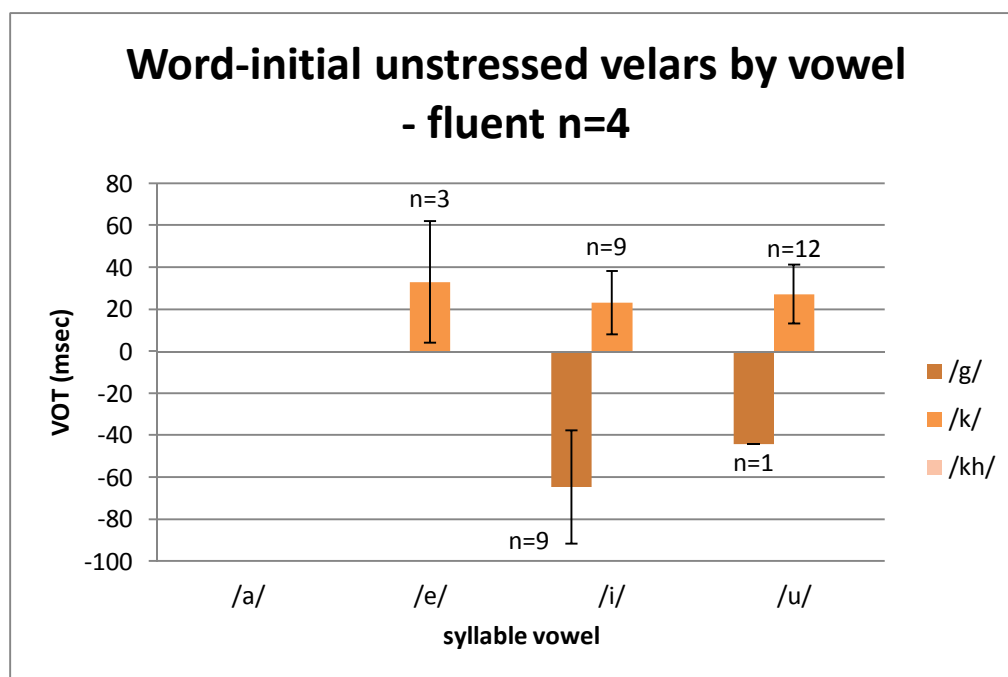


Figure 69.

Talkers A-D (fluent): mean voice onset time for word initial unstressed velars by individual vowel; the number of tokens contributing to each mean is presented above/below the bar. Missing bars indicate that the talker did not produce any usable tokens for that target consonant.

Word initial stressed stops by vowel:

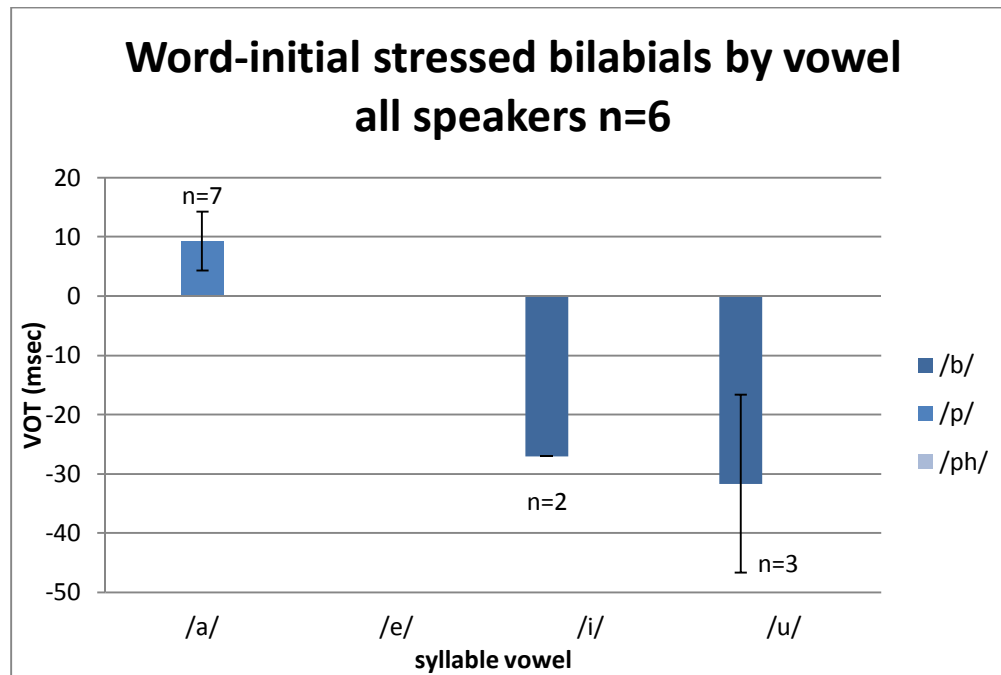


Figure 70.

Talkers A-F (all talkers): mean voice onset time for word initial stressed bilabials by individual vowel; the number of tokens contributing to each mean is presented above/below the bar. Missing bars indicate that the talker did not produce any usable tokens for that target consonant.

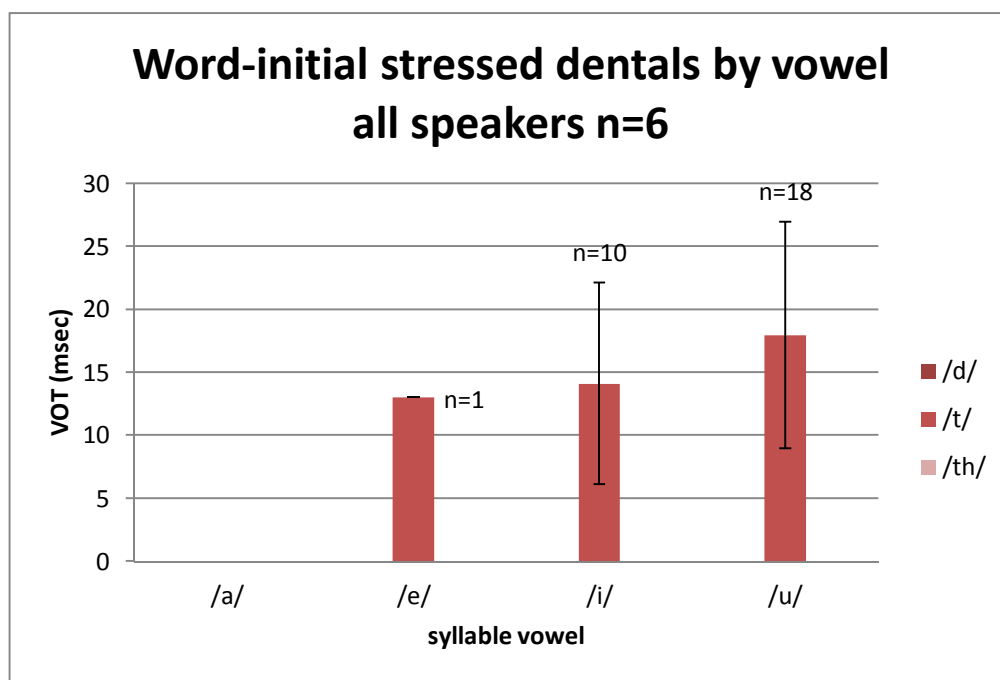


Figure 71.

Talkers A-F (all talkers): mean voice onset time for word initial stressed dentals by individual vowel; the number of tokens contributing to each mean is presented above/below the bar. Missing bars indicate that the talker did not produce any usable tokens for that target consonant.

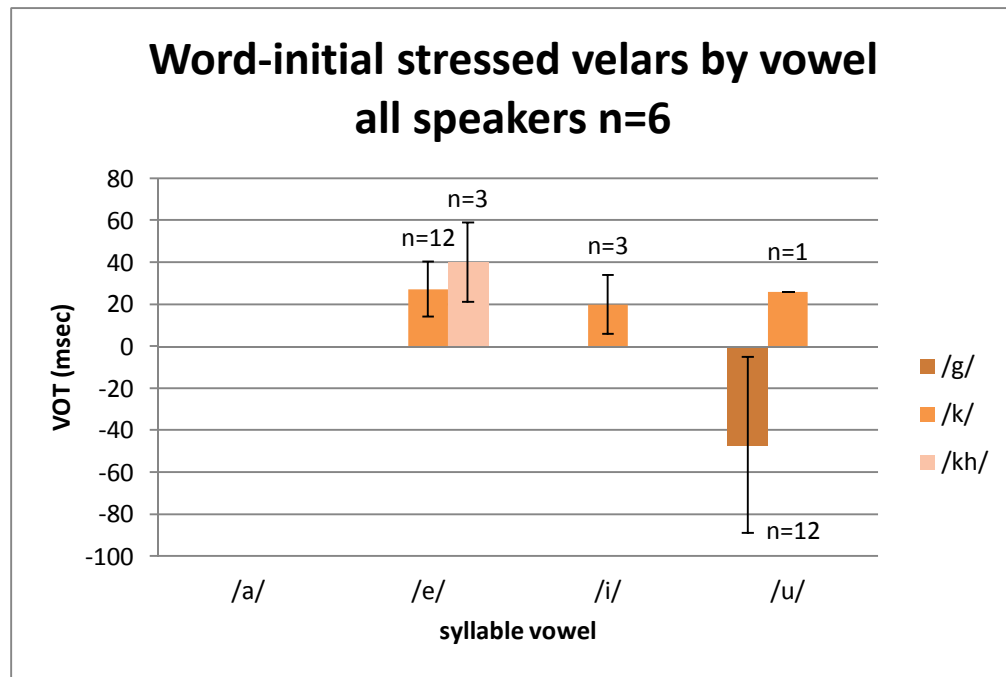


Figure 72.

Talkers A-F (all talkers): mean voice onset time for word initial stressed velars by individual vowel; the number of tokens contributing to each mean is presented above/below the bar. Missing bars indicate that the talker did not produce any usable tokens for that target consonant.

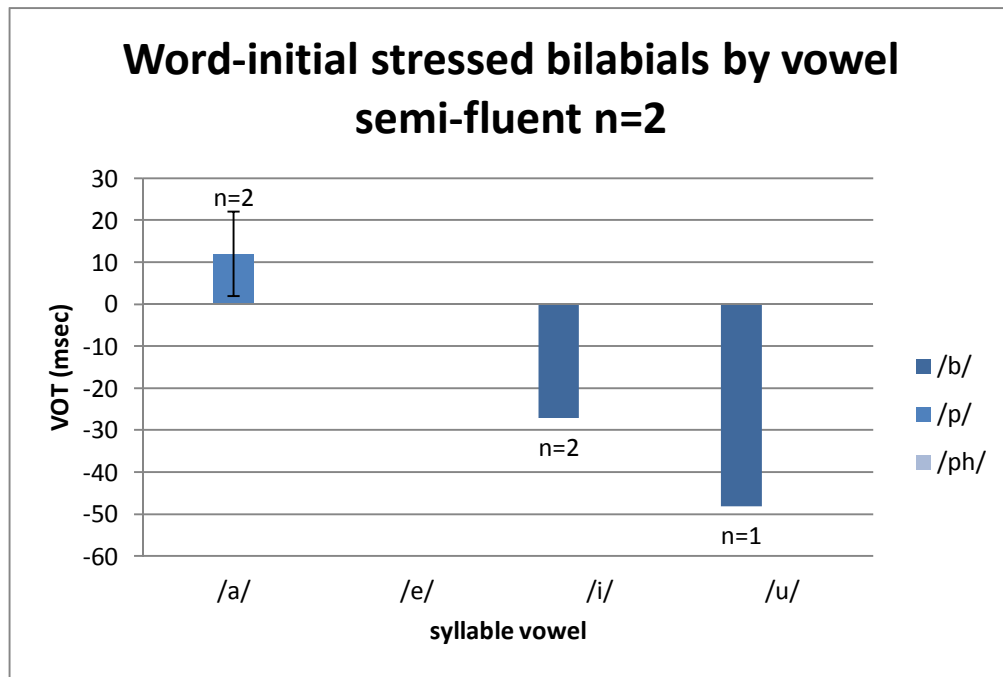


Figure 73.

Talkers E and F (semifluent): mean voice onset time for word initial stressed bilabials by individual vowel; the number of tokens contributing to each mean is presented above/below the bar. Missing bars indicate that the talker did not produce any usable tokens for that target consonant.

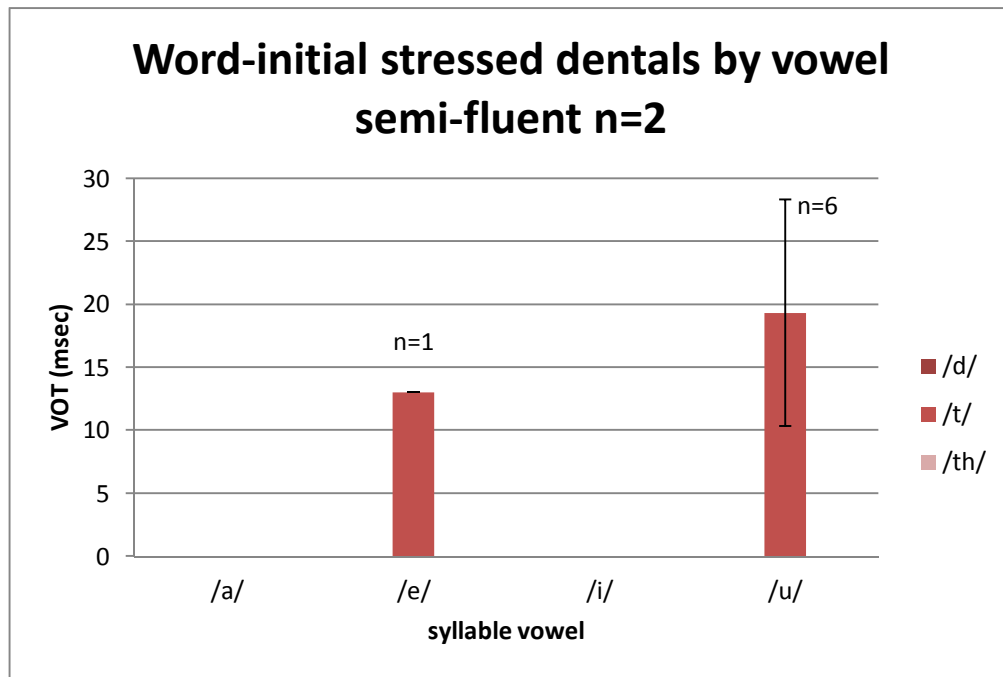


Figure 74.

Talkers E and F (semifluent): mean voice onset time for word initial stressed dentals by individual vowel; the number of tokens contributing to each mean is presented above/below the bar. Missing bars indicate that the talker did not produce any usable tokens for that target consonant.

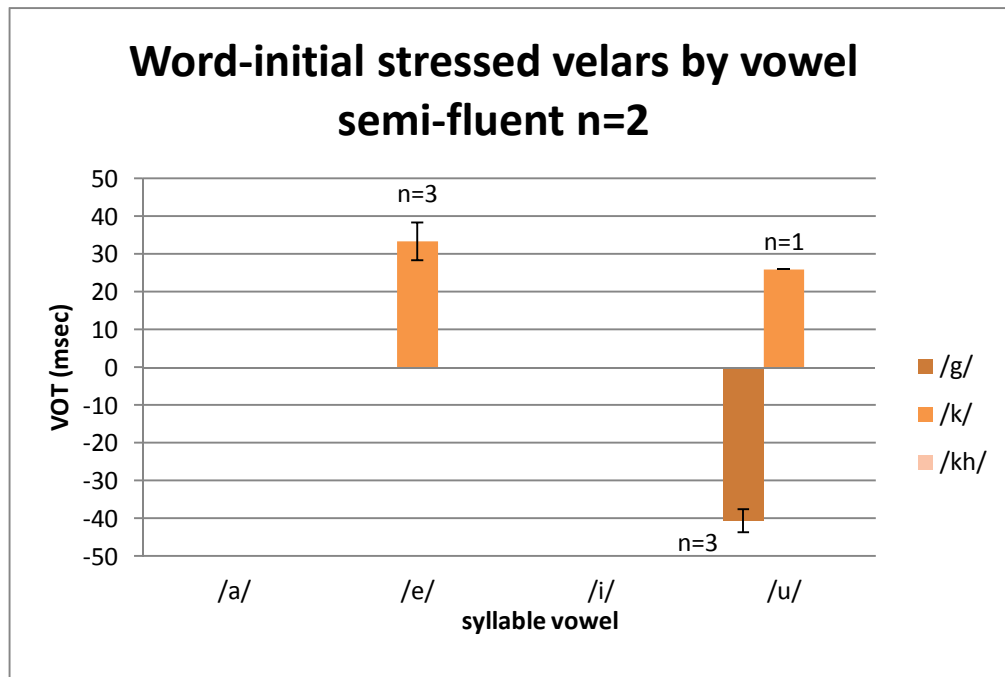


Figure 75.

Talkers E and F (semifluent): mean voice onset time for word initial stressed velars by individual vowel; the number of tokens contributing to each mean is presented above/below the bar. Missing bars indicate that the talker did not produce any usable tokens for that target consonant.

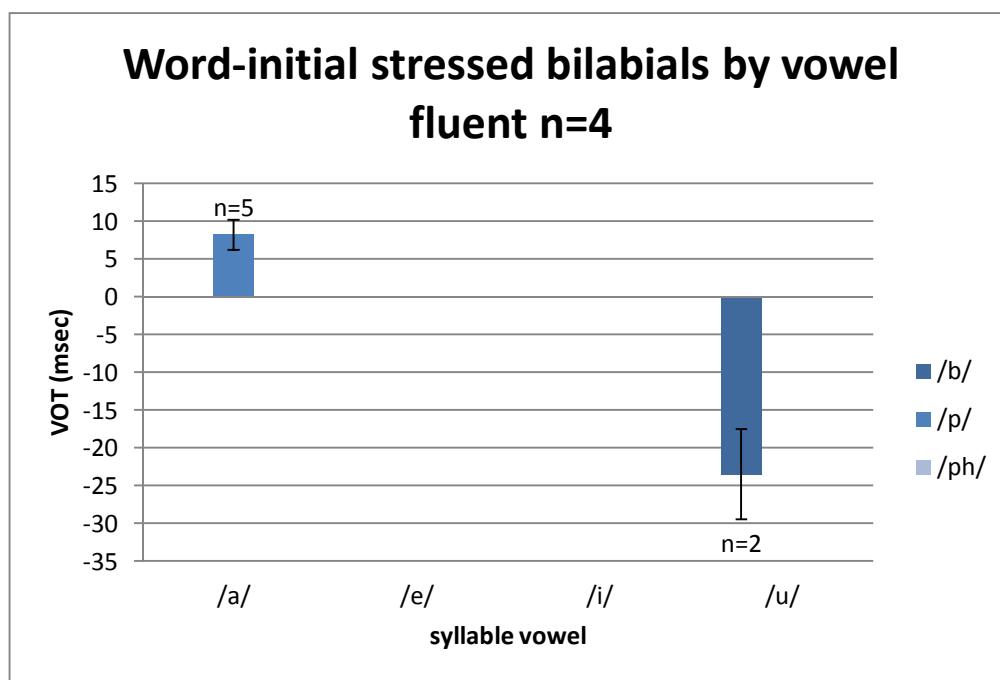


Figure 76.

Talkers A-D (fluent): mean voice onset time for word initial stressed bilabials by individual vowel; the number of tokens contributing to each mean is presented above/below the bar. Missing bars indicate that the talker did not produce any usable tokens for that target consonant.



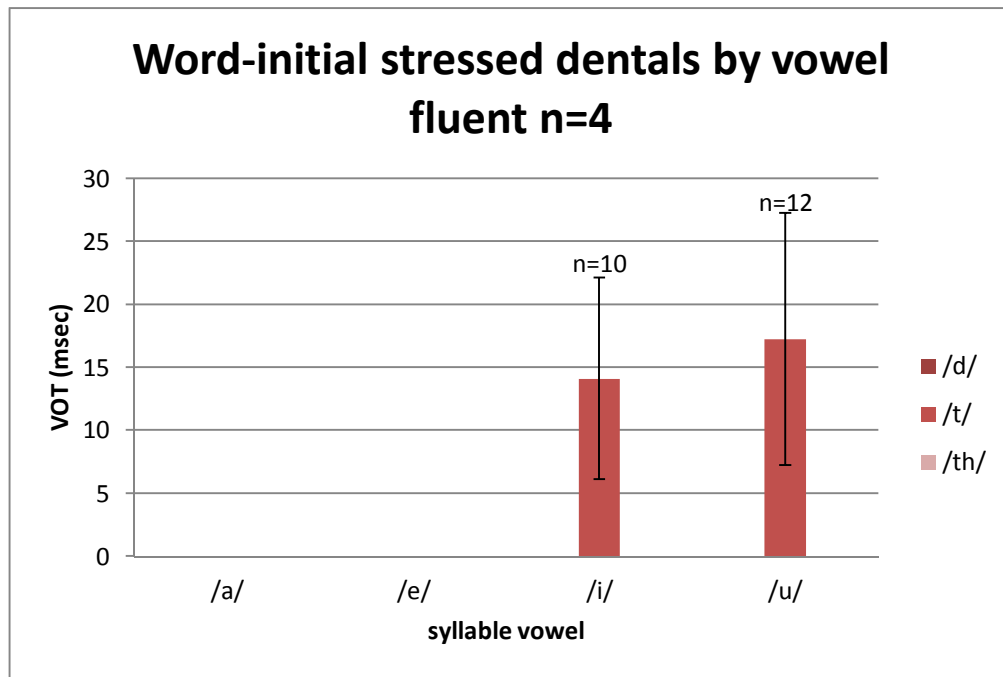


Figure 77.

Talkers A-D (fluent): mean voice onset time for word initial stressed dentals by individual vowel; the number of tokens contributing to each mean is presented above/below the bar. Missing bars indicate that the talker did not produce any usable tokens for that target consonant.

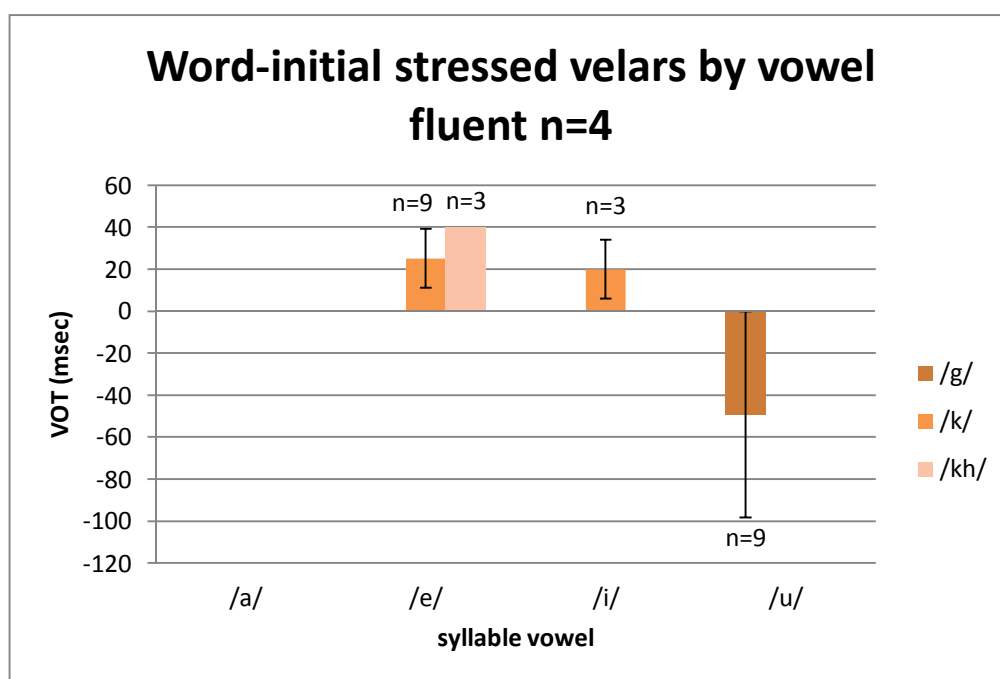


Figure 78.

Talkers A-D (fluent): mean voice onset time for word initial stressed velars by individual vowel; the number of tokens contributing to each mean is presented above/below the bar. Missing bars indicate that the talker did not produce any usable tokens for that target consonant.

Word medial unstressed stops by vowel:

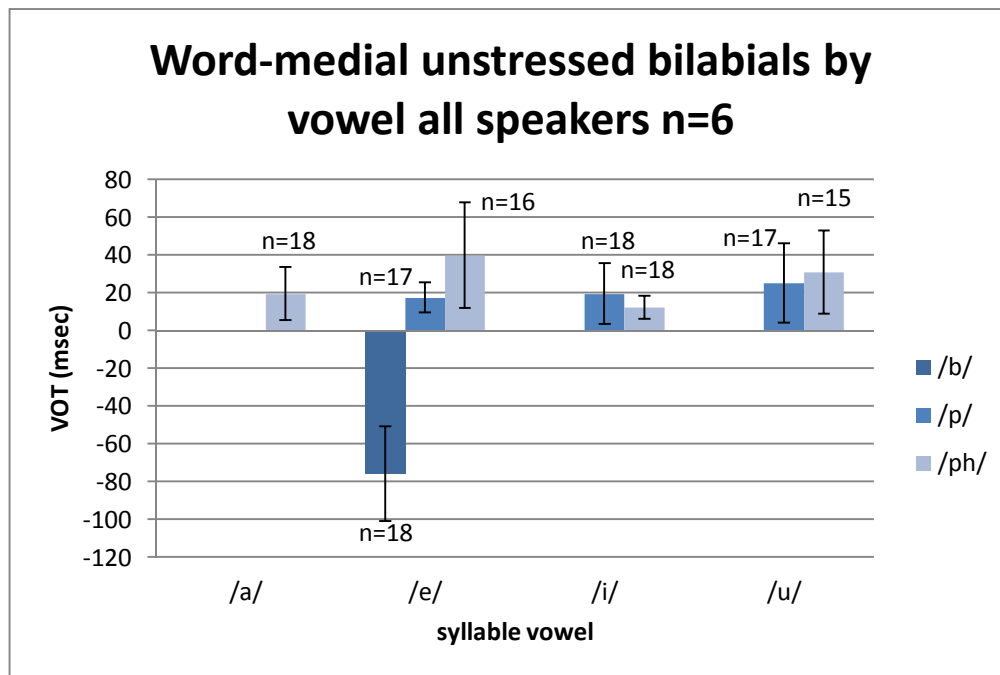


Figure 79.

Talkers A-F (all talkers): mean voice onset time for word medial unstressed bilabials by individual vowel; the number of tokens contributing to each mean is presented above/below the bar. Missing bars indicate that the talker did not produce any usable tokens for that target consonant.

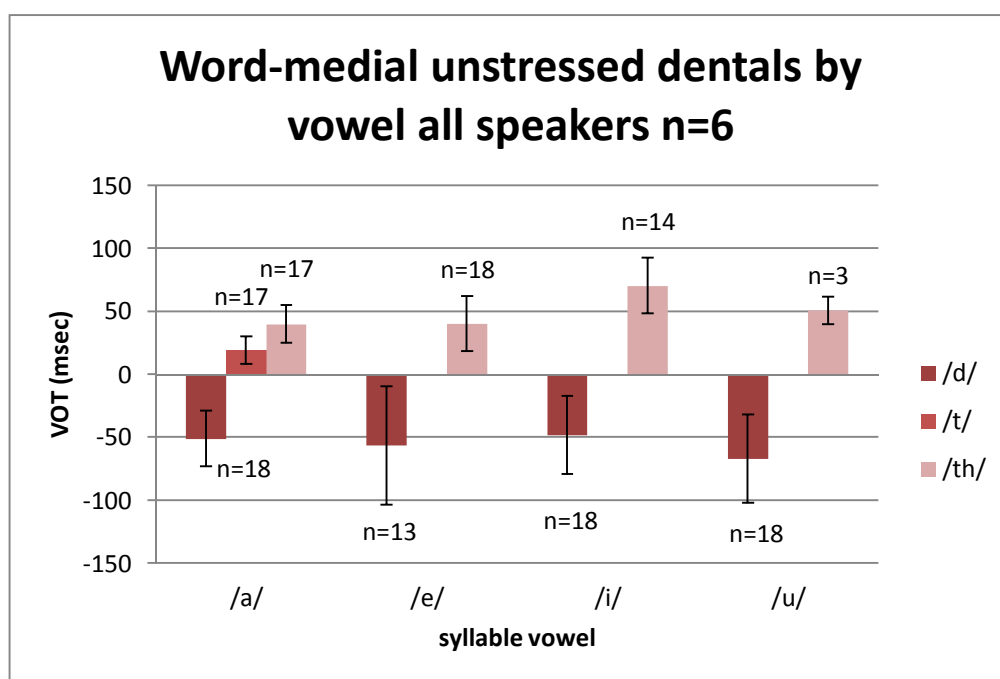


Figure 80.

Talkers A-F (all talkers): mean voice onset time for word medial unstressed dentals by individual vowel; the number of tokens contributing to each mean is presented above/below the bar. Missing bars indicate that the talker did not produce any usable tokens for that target consonant.

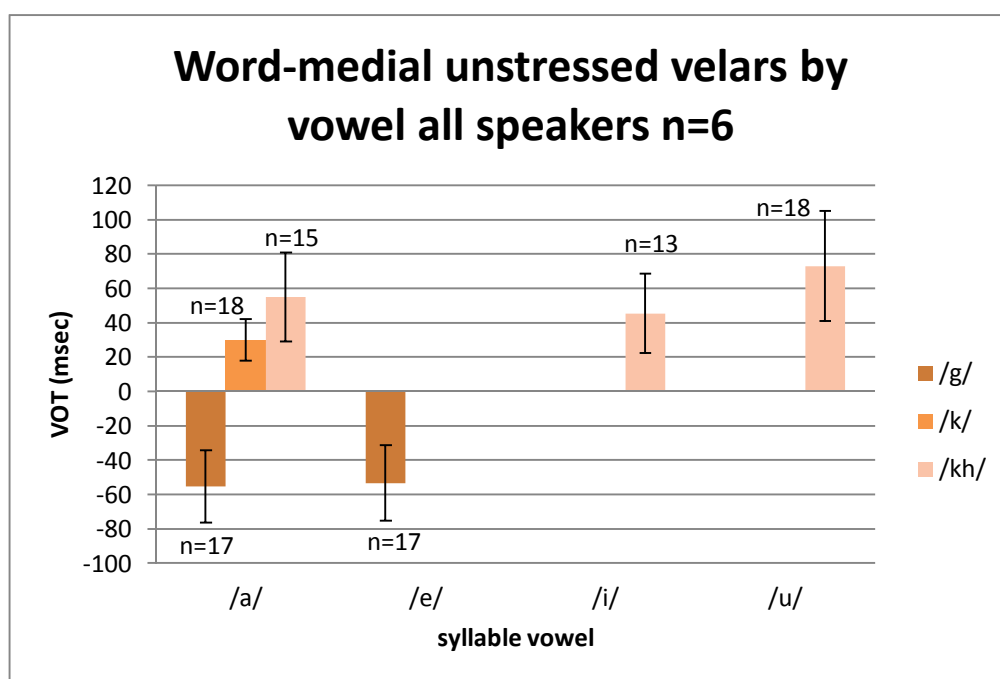


Figure 81.

Talkers A-F (all talkers): mean voice onset time for word medial unstressed velars by individual vowel; the number of tokens contributing to each mean is presented above/below the bar. Missing bars indicate that the talker did not produce any usable tokens for that target consonant.

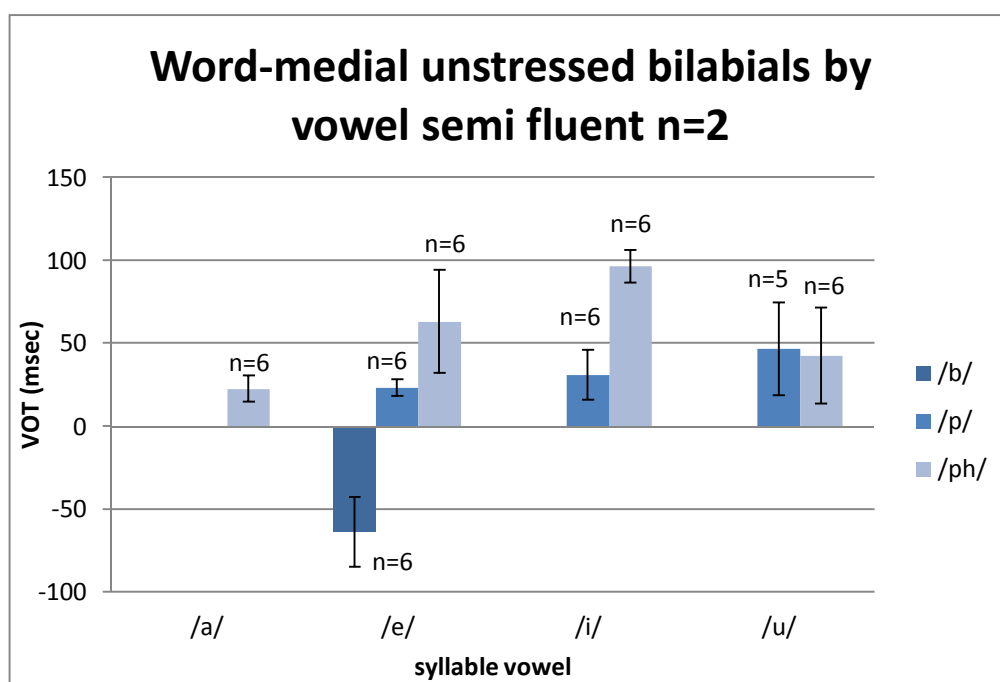


Figure 82.

Talkers E and F (semifluent): mean voice onset time for word medial unstressed bilabials by individual vowel; the number of tokens contributing to each mean is presented above/below the bar. Missing bars indicate that the talker did not produce any usable tokens for that target consonant.

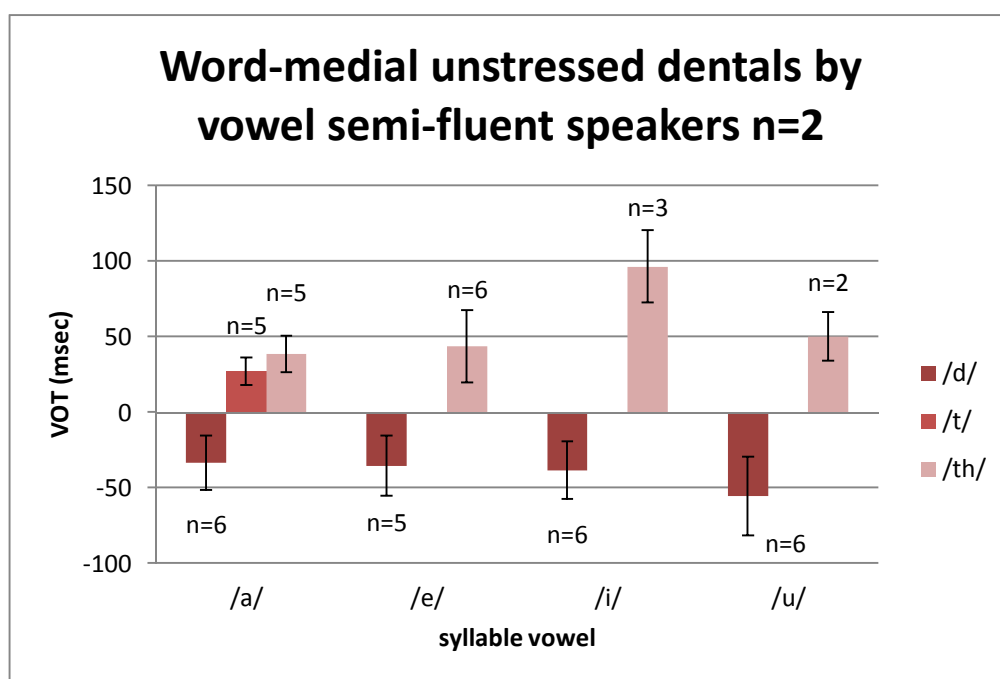


Figure 83.

Talkers E and F (semifluent): mean voice onset time for word medial unstressed dentals by individual vowel; the number of tokens contributing to each mean is presented above/below the bar. Missing bars indicate that the talker did not produce any usable tokens for that target consonant.

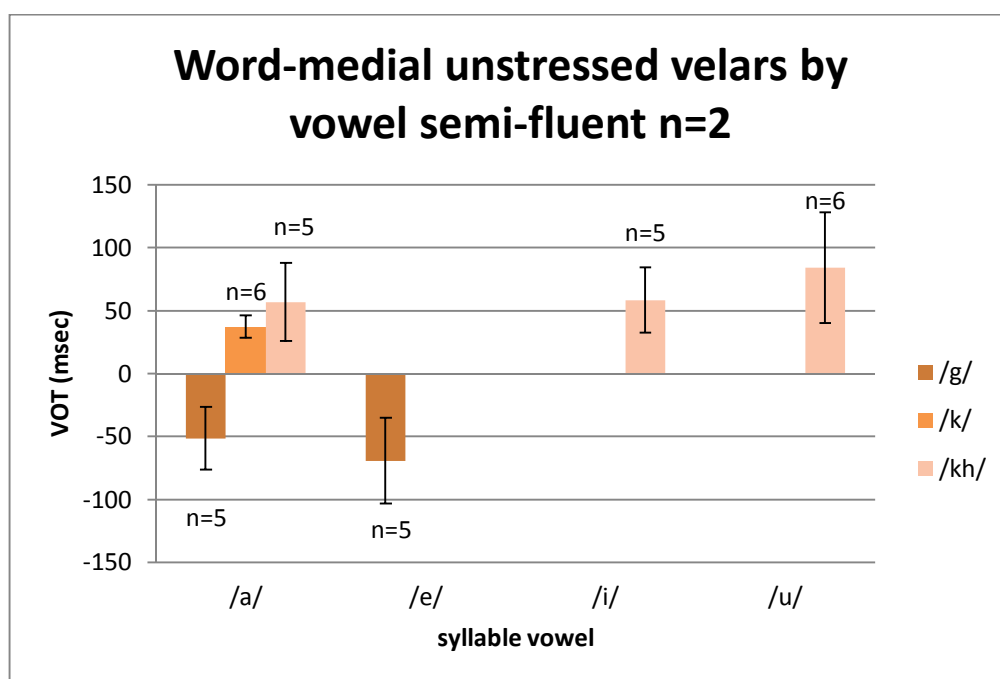


Figure 84.

Talkers E and F (semifluent): mean voice onset time for word medial unstressed velars by individual vowel; the number of tokens contributing to each mean is presented above/below the bar. Missing bars indicate that the talker did not produce any usable tokens for that target consonant.



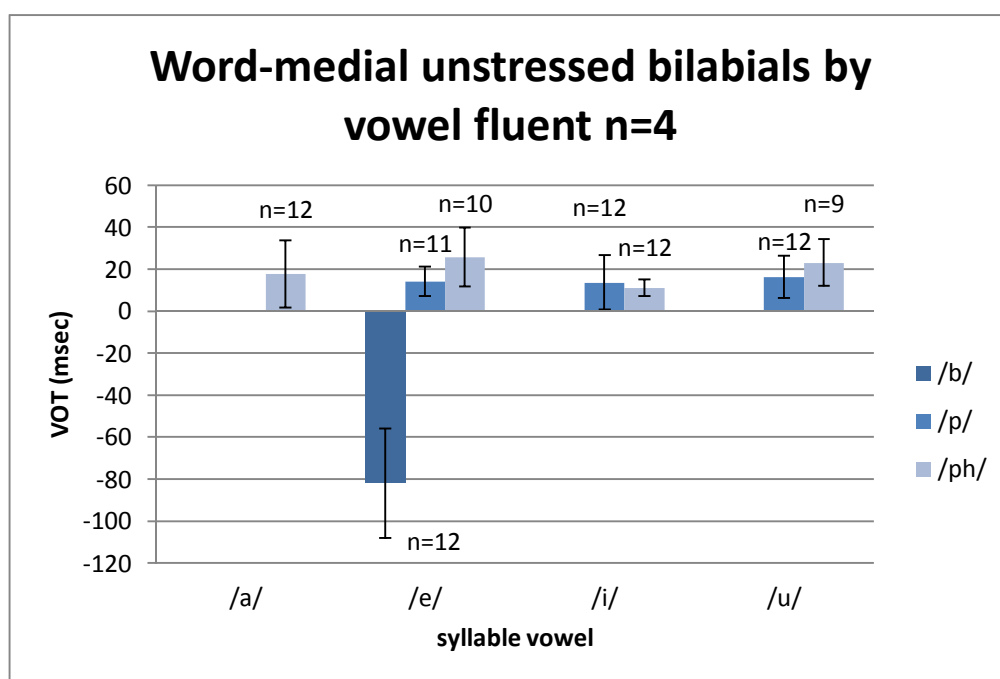


Figure 85.

Talkers A-D (fluent): mean voice onset time for word medial unstressed bilabials by individual vowel; the number of tokens contributing to each mean is presented above/below the bar. Missing bars indicate that the talker did not produce any usable tokens for that target consonant.

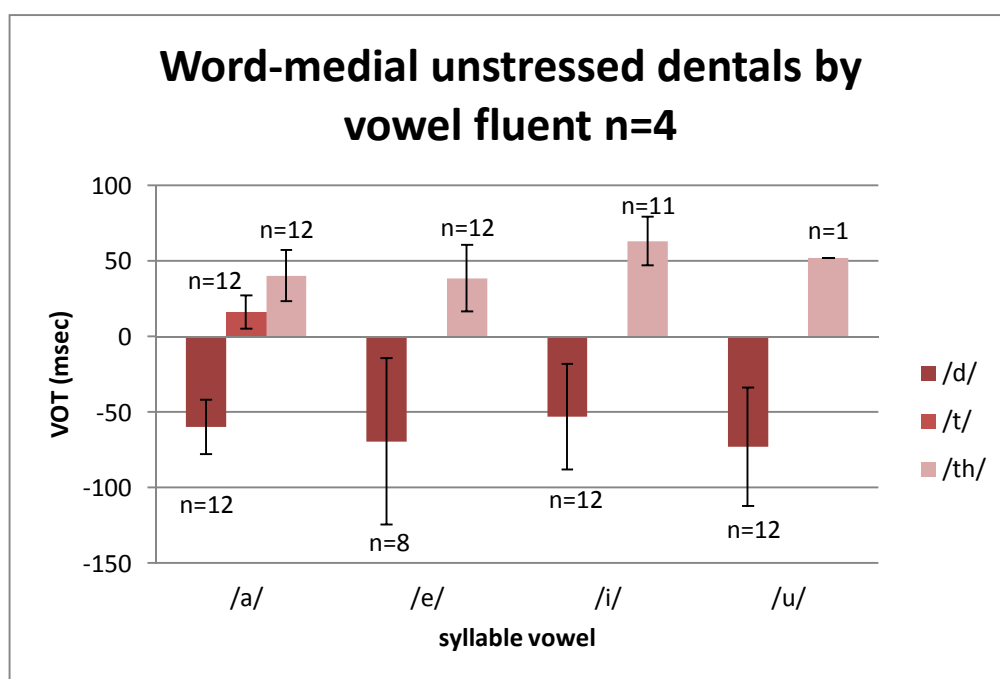


Figure 86.

Talkers A-D (fluent): mean voice onset time for word medial unstressed dentals by individual vowel; the number of tokens contributing to each mean is presented above/below the bar. Missing bars indicate that the talker did not produce any usable tokens for that target consonant.

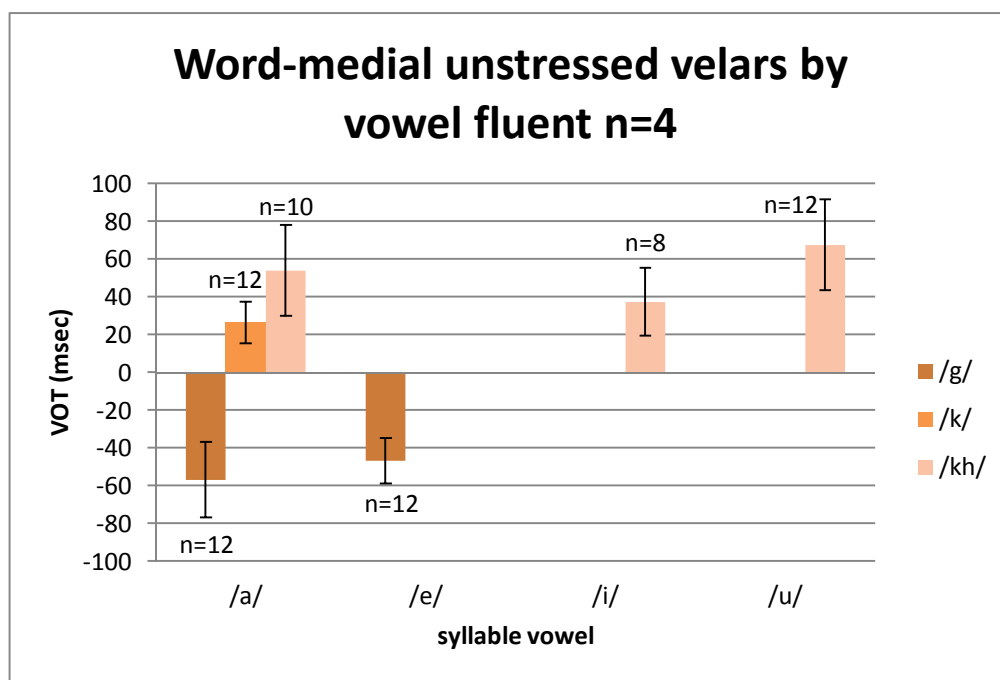


Figure 87.

Talkers A-D (fluent): mean voice onset time for word medial unstressed velars by individual vowel; the number of tokens contributing to each mean is presented above/below the bar. Missing bars indicate that the talker did not produce any usable tokens for that target consonant.

Word medial stressed stops by vowel:

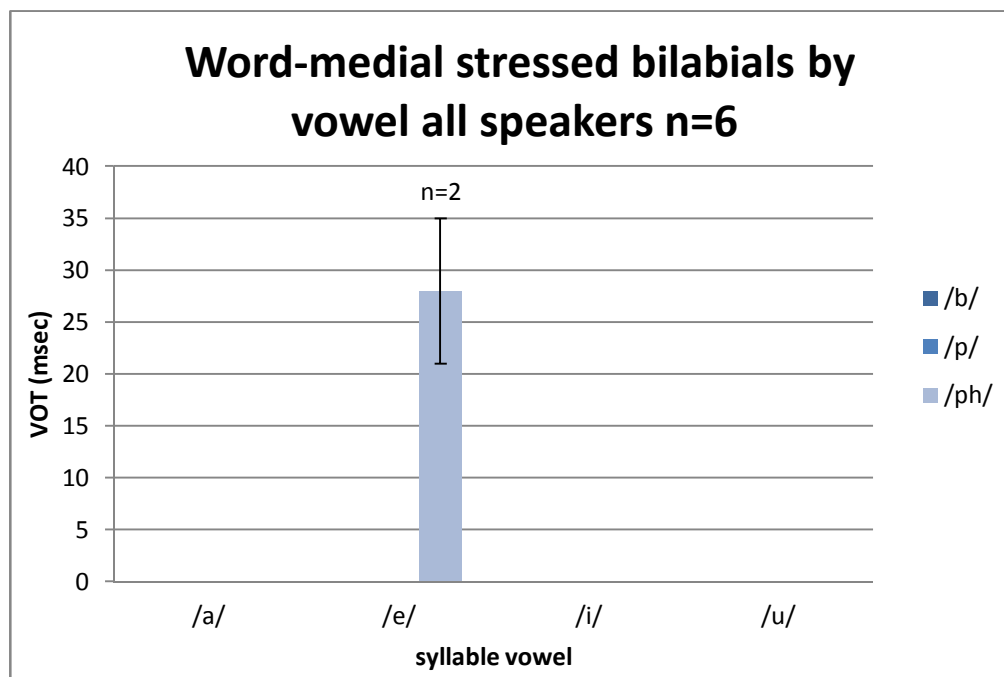


Figure 88.

Talkers A-F (all talkers): mean voice onset time for word medial stressed bilabials by individual vowel; the number of tokens contributing to each mean is presented above/below the bar. Missing bars indicate that the talker did not produce any usable tokens for that target consonant.

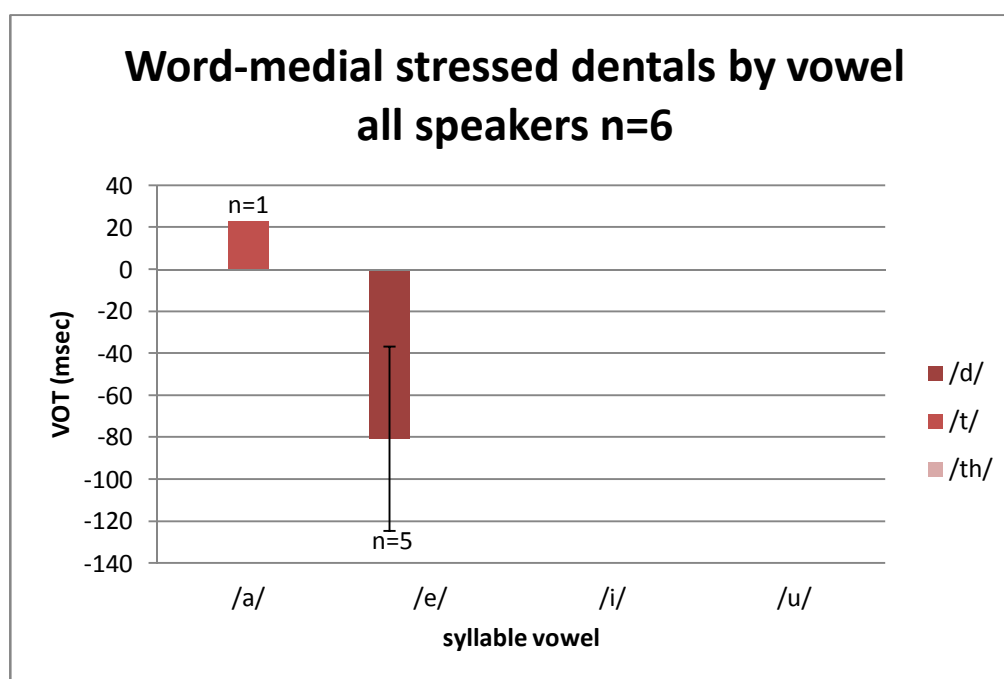


Figure 89.

Talkers A-F (all talkers): mean voice onset time for word medial stressed dentals by individual vowel; the number of tokens contributing to each mean is presented above/below the bar. Missing bars indicate that the talker did not produce any usable tokens for that target consonant.

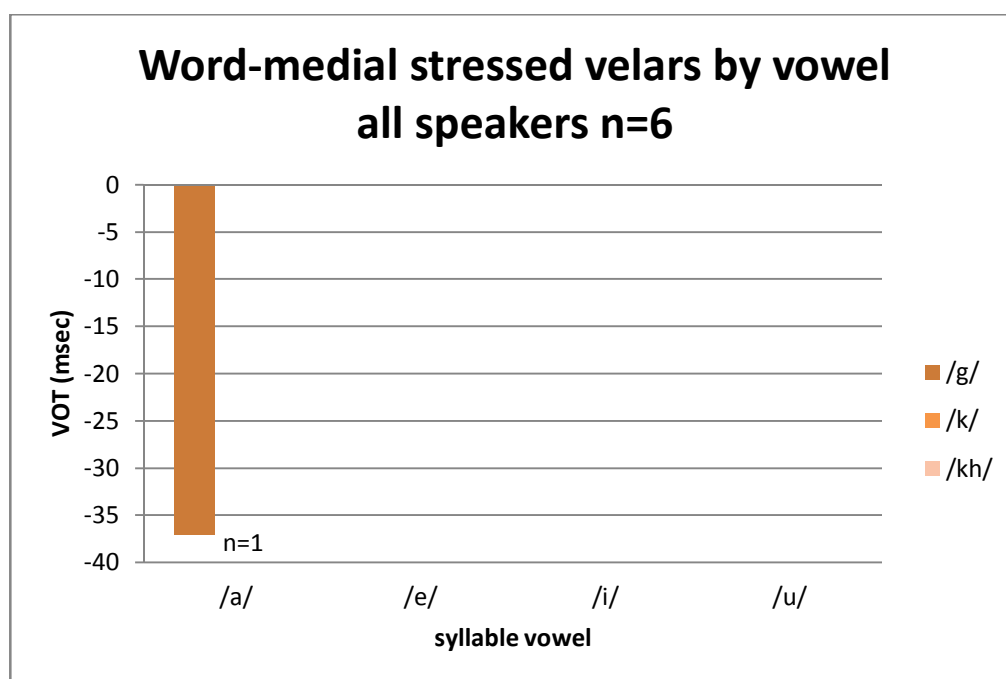


Figure 90.

Talkers A-F (all talkers): mean voice onset time for word medial stressed velars by individual vowel; the number of tokens contributing to each mean is presented above/below the bar. Missing bars indicate that the talker did not produce any usable tokens for that target consonant.

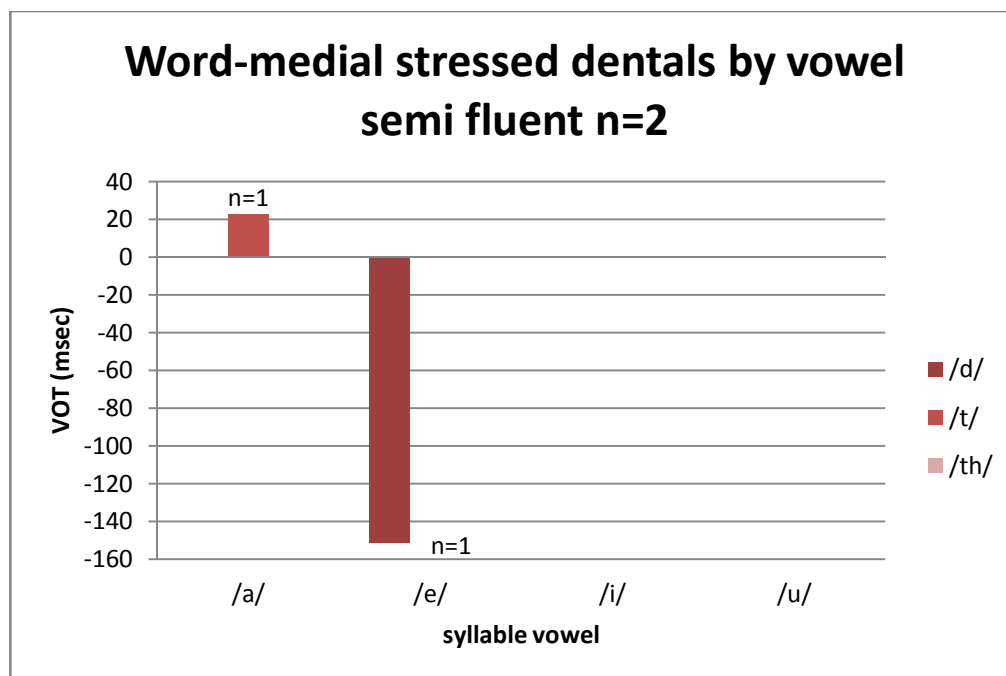


Figure 91.

Talkers E and F (semifluent): mean voice onset time for word medial stressed dentals by individual vowel; the number of tokens contributing to each mean is presented above/below the bar. Missing bars indicate that the talker did not produce any usable tokens for that target consonant.

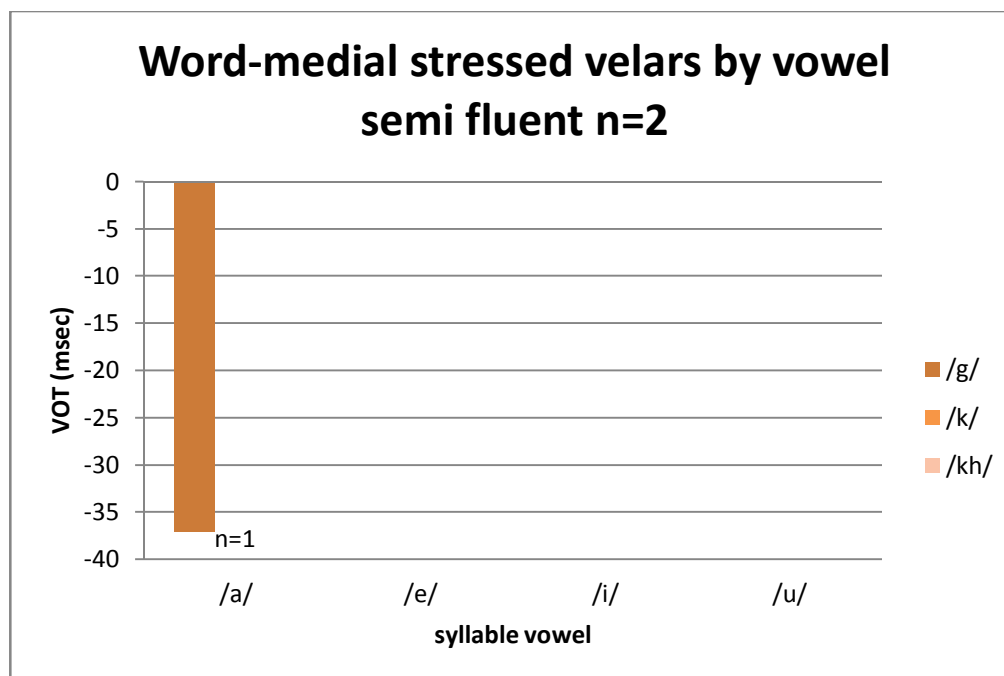


Figure 92.

Talkers E and F (semifluent): mean voice onset time for word medial stressed velars by individual vowel; the number of tokens contributing to each mean is presented above/below the bar. Missing bars indicate that the talker did not produce any usable tokens for that target consonant.



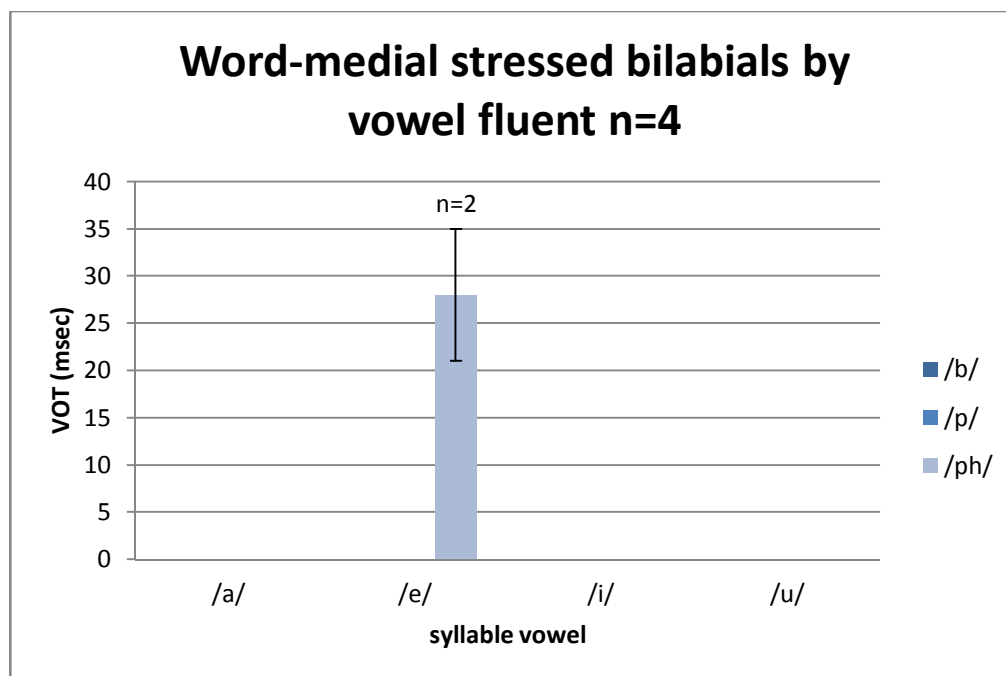


Figure 93.

Talkers A-D (fluent): mean voice onset time for word medial stressed bilabials by individual vowel; the number of tokens contributing to each mean is presented above/below the bar. Missing bars indicate that the talker did not produce any usable tokens for that target consonant.

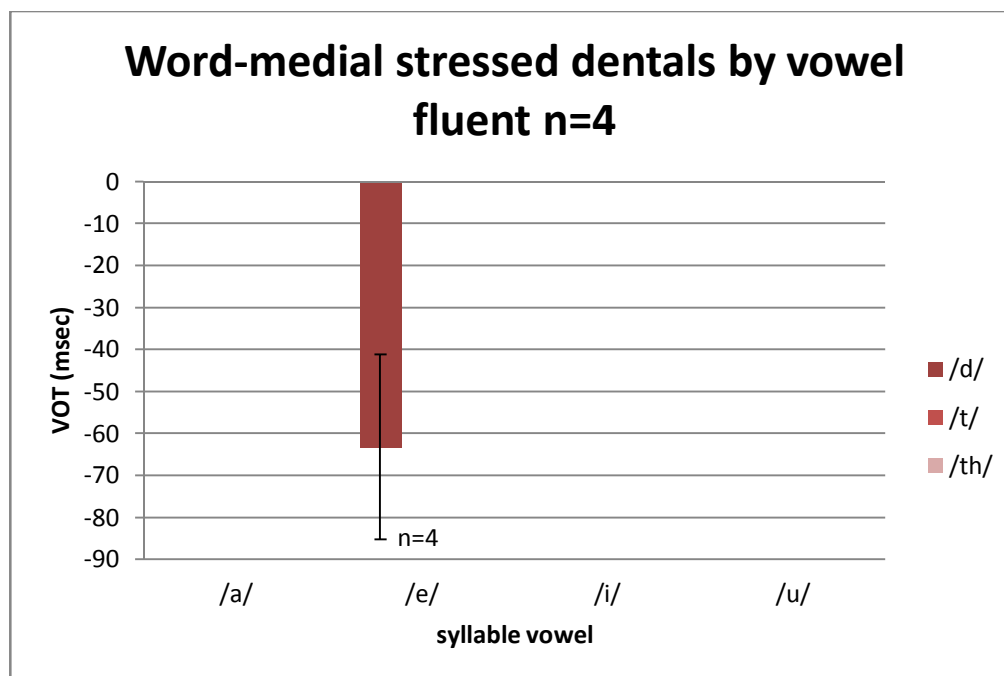


Figure 94.

Talkers A-D (fluent): mean voice onset time for word medial stressed dentals by individual vowel; the number of tokens contributing to each mean is presented above/below the bar. Missing bars indicate that the talker did not produce any usable tokens for that target consonant.

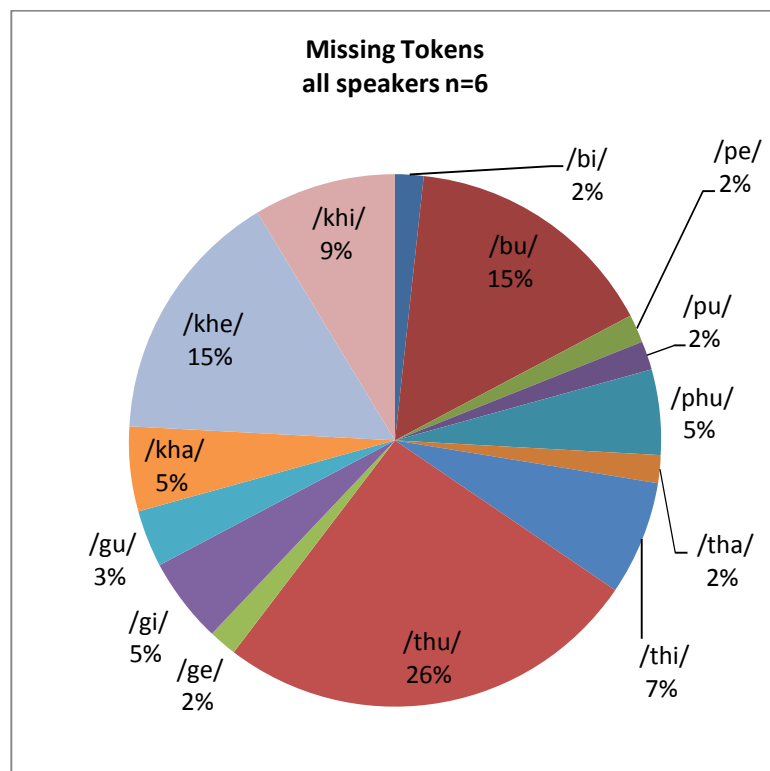


Figure 95.

Missing tokens by target consonant. Missing tokens are those that either were not elicited or were not usable in the data.

## V. DISCUSSION

Based on results from this study, the original hypothesis is predominantly supported, but results also reveal some ambiguity where the hypothesized pattern was not expressed. There appear to be clear distinctions between three manners of oral stop in Omaha, but certain analyses show exceptions to this generalization. The millisecond values associated with the three categories are consistent with Lisker and Abramson's (1964) voiced, voiceless unaspirated (short lag), and voiceless aspirated (long lag) categories. Most of the acoustic analysis supports that the stop categories in Omaha are those listed above, rather than being, for example, "partially aspirated" as proposed of one Korean category, or "highly aspirated" (110+ msec) as proposed of Navajo and Tlingit aspiration (Cho and Ladefoged 1999).

**Error! Reference source not found.** specifies which charts from the current study support the hypothesis, which do not support it, and which are inconclusive. The criteria for the first category in the individual tokens chart is any chart where at least nine tokens from each category are distinct from those of a neighboring category in a way that reflects the hypothesized three way stop contrast. For a bar graph to meet the criteria for supporting the hypothesis, it needs to show a three way pattern in every case where enough data is available for three way contrasts to be measured. The criterion for the "do not support" (**Error! Reference source not found.**) category is: any evidence in bar

graphs that the three way contrast is not borne out by data where enough data are available for comparison. Finally, the -inconclusiveøcategory comprises

119

all graphs which do not contain enough data for even one three way contrast to be visible given the constraints of the graph.

Forty-four charts support the hypothesis, 31 are inconclusive, and 19 do not support the hypothesis. There is a disproportionate number of individual VOT token charts represented in the -do not supportøcategory. This is because there were very frequent cases where aspirated and unaspirated tokens had very similar msec values. This is the principal section of the experiment that does not support the hypothesis. Averages across analyses support the hypothesis, except where there are too few tokens to show all three stop manners given the criteria of the cross analysis.

Additionally, the robustly supported theory that VOT is longer for farther back places of articulation is also supported with data from the current study. This is especially visible in individual- and cross-speaker averages (see Figure 20, Figure 21; compare values from Figure 85 through Figure 87). These five figures illustrate average VOTs that increase in a pattern of bilabial < dental < velar.

To further explain how the data from this study supports the hypothesis that Omaha contains a voiced-unaspirated-aspirated contrast, a closer comparison of the Lisker and Abramson (1968) data to these data are offered. In that study:

- Stops in isolated words in two-, three-, and four-category systems were examined in the first analysis.

- Then, stops in running speech were examined, both in initial and noninitial positions.

120

- Thai and Armenian (three category languages) lined up well with two-category languages: msec values: -100, +10, +75 averages for voiced, voiceless unaspirated, and voiceless aspirated. Korean is peculiar in that all stops are in the positive half of the continuum.
- Sentence data are congruent with word data (p. 413).
- Sentence embedded stops showed unbroken voicing for voiced stops and in English unaspirated stops. But all other languages: unaspirated stops did show broken voicing even in sentences.

Lisker and Abramson acknowledge that although they controlled for word position (they specify word initial and nonword initial positions), they did not control for all variables (e.g. rate of speech, stress, vocalic environment, and allophonic variants of phonemes). The current study reports data both in general, without elimination of confounds, and in specific, with stress, word position and vowel environment isolated. However, only word position analyses can be compared directly with Lisker and Abramson's findings, which were analyzed in a parallel way. Lisker and Abramson's findings for isolated words are not compared here because no such tests were conducted in this experiment.

The relevant section uses data from sentences for word initial and non-word initial stops followed by vowels in languages with two- and three-category stop systems.

**Error! Reference source not found.** is a summary of data (created by the author for

the current paper) from Lisker and Abramson (1964). It shows average values for stops in sentences from languages with a three way contrast.

121

The corresponding average values from Omaha would be those analyzed by word position only. Because this data does not control for the syllabic vowel used in the target syllable, data containing all four vowel /a/, /e/, /i/, and /u/ and combined below. In most cases, the four sentences representing these vowels included a combination of word initial and word medial syllables. In the cases of /p<sup>h</sup>/, /d/, and /t<sup>h</sup>/, all four syllables were word medial, leaving a gap in the data in the word initial part of the chart. Means are as follows, and include fluent and semifluent speakers (**Error! Reference source not found.**):

The VOT averages of most stops across places of articulation in Omaha are within 5 to 15 msec of averages for the corresponding category and place of articulation across Armenian, Thai, and Korean. The exceptions to this are /p<sup>h</sup>/, /k<sup>h</sup>/, /d/, and /g/. /p<sup>h</sup>/ has a shorter msec average in both word positions in Omaha than in any averages from Lisker and Abramson's word position analyses. This could be due to the accidental incorporation of a possibly unaspirated stop to the stimulus item for aspirated /p<sup>h</sup>/ (discussed below). /k<sup>h</sup>/ has aspiration that is more similar to the -slightly aspirated category proposed by Cho and Ladefoged (1999) and others, with a VOT average of 40 msec. This could mean that aspirated stops in Omaha are actually -slightly aspirated but further study would be required to confirm this, as a number of stops also showed higher VOT values.

Based on the discussion to this point, there is reasonable evidence in support of the conclusion that Omaha has a three way stop contrast comprising voiced, voiceless unaspirated, and voiceless aspirated categories. It must be noted, however, that this study implicitly assumes written transcripts of Omaha from 1890 and present day Omaha

122

recordings to be representing the same speech sounds. It is possible that pronunciation has evolved in Omaha since 1890, due to attrition in general and English influence in particular, but language change likely has not been so extreme as to render the three categories unrecognizable today compared to the Dorsey transcripts. Originally it was hypothesized that negative transfer from Dorsey's English background could have influenced his transcripts. This could be true, but is impossible to determine in this study. It is only possible to draw conclusions on the language as it is spoken today, and not possible to surmise whether the phonetic qualities of the assumed three way system were the same in Dorsey's time as they are today. This means that if we find discrepancies between the msec values recorded in the current study and the phonetic characters in Dorsey's transcripts, it may or may not be due to inaccurate transcription on Dorsey's part. Ultimately, recommendations made about pronunciation based on this study are situated within the scope of Omaha elders' lifetimes and the pronunciation they are personally familiar with, rather than the pronunciation of their grandparents or great-grandparents.

Another component of the hypothesis was that based on cross language data (Maddieson 1984), aspirated voiceless stops occur less frequently than unaspirated voiceless stops. It was hypothesized that if Dorsey's transcripts had been shown to



contain flaws, it would be in the transcription of too many aspirates compared with those that really exist in Omaha. However, this effect was also difficult to address here for the reasons described above. Additionally, because only one word was used to represent each of the nine stops combined with each of four vowels, it is impossible to determine

123

whether, if there is a discrepancy between this study and Dorsey's transcriptions, it is due to one word changing over time, a larger scale sound change, negative transfer in Dorsey's idiolect, an idiosyncrasy in the idiolect of the participant speaker in this study, or a true lowered incidence of aspirated stops versus unaspirated ones. This is also true of the missing tokens: we could hypothesize that those stops which were most difficult to elicit were also in the least frequently occurring series. But the format of this study makes this impossible to determine. The missing tokens could also be due to an individual word being antiquated, irrelevant, or rare.

Although data on contemporary Omaha appear fairly conclusive, there are a number of limitations present in this study that will be discussed below. Some limitations are due to the design of this study and others are inherent to the situation of an endangered language.

Limitations of the experimental design are listed here. The first limitation is of mistaken identity stimulus tokens. For instance, sometimes one manner of stop seems to contain values traditionally associated with another manner. In these cases it is hypothesized that our assumptions about the word's stop manner were incorrect to begin with. This is supported by data points as shown in figures for "All VOT tokens by speaker". The first three data points in each chart show the three repetitions from each

speaker of stops preceding /a/; the next three points represent stops preceding /e/, etc. We can see in the bilabials charts that the two most fluent speakers, A and B, have prototypically unaspirated values for the supposedly aspirated stimulus words. Rather than considering this to be aberrant, it is likely to be an example of a stimulus word that

124

was incorrectly categorized. The word, *zháp<sup>h</sup>ahi*, means 'rosin weed'. However, when working with the Omaha elder who aided in writing the sentences and verifying grammaticality, it was learned that the word these days means 'Indian gum' or just 'chewing gum' since rosin weed was a plant that Omaha people used for chewing. *zháp<sup>h</sup>ahi* was thus glossed henceforth as 'Indian gum'. Most participants shortened the word to *zháp<sup>h</sup>a*. Unfortunately for the purposes of this study, shortening the word also changes the aspiration of /p<sup>h</sup>a/ to /pa/ (Eschenberg, personal communication). Thus, the first three green tokens in the bilabials charts show three repetitions of *zháp<sup>h</sup>a*, many of which could have been intentionally uttered without aspiration, as is purportedly grammatical. It is in this way that we can hypothesize that certain supposed categories represented in specific stimulus words could be misplaced, skewing the results.

Next, this study does not measure the same syllable in multiple environments, i.e., both as isolated words and in running speech. Neither does the study examine the same syllable both word initially and word medially. If either of these controls had been added to the experiment, influences from word position and from running speech could have been more easily accounted for. Another way to focus on sentence level influences would have been to isolate the local vowel environments. Rather than testing each stop with each of the four vowels, the study could have tested uniformly each stop with only one

vowel following in the syllable across all stops in consideration. To this we could have added more words. For example, if an experiment occurred in which all stops were followed by /a/, then each stop /ba/, /pa/, /p<sup>h</sup>a/, /da/, /ta/, /t<sup>h</sup>a/, /ga/, /ka/, and /k<sup>h</sup>a/ could have been tested in a variety of words. This would yield results for more than one vocalic

125

environment. Thus, generalizations about VOTs would truly be founded in a more general test of the stops, rather than a test of merely one word per stop manner and vowel.

An in depth study of the allophonic variation of sounds in the words chosen for stimulus items was not undertaken. Allophonic variation could contribute to the VOT data in ways not predicted by this study. Finally, some standard deviation values were large, possibly due to confounds discussed above. This does detract from decisiveness of the conclusions. Nevertheless, the values are aligned closely enough to those in Lisker and Abramson that it is still valid to assert the voiced, voiceless-unaspirated, and voiceless-aspirated contrast discussed herein.

The administration of the stimulus items was also not controlled. This was largely beyond the control of this study, since it related to the degree of fluency of each speaker, of which few remain today. It was intended that English sentences would be read aloud to the participants, who would repeat in Omaha. However, some participants preferred to read the list, which included English and Fletcher-LaFlèche orthography Omaha. In total, four participants read the list themselves, and two had the list read to them. This could perhaps contribute to discrepancies in the data.

Of course, other variables were also impossible to control due to the endangered status of Omaha. The most prominent of these was the small number of Omaha fluent speakers alive today, only a fraction of whom could be contacted. This yielded a less than ideal sample size. Indeed, the study included two semifluent speakers to increase sample size, but ideally only fully fluent speakers would have been included in the study. In light

126

of the small number of speakers, it was unrealistic to exclude participants based on hearing or speech impairments. As a result, at least two participants were both hearing- and speech-impaired. Nevertheless, it should be highlighted that considering how few speakers of Omaha remain today, an *n* of 6 is a large percentage of the absolute total possible.

#### Future research

There is a great deal of additional research that could be done in this vein. Based on Miller (2001), listeners can identify better and worse exemplars of individual phonemes in their native language. Such a study could be done working with native Omaha speaking participants to ascertain category goodness ratings in the nine sounds analyzed phonetically in the current study. This could shed additional light on which msec values are ideal for a given stop at the phonemic level.

Following Cho and McQueen (2005), who suggest stress influences VOT, work could be done on the influence of stress on VOT in Omaha. Siouan expert Robert Rankin also suggests undertaking an analysis of Omaha stress patterns (personal

communication). Based on certain minimal pairs it seems likely that stress is phonemic in Omaha, although a formal study has not been conducted.

Also, the present experiment could be repeated with additional stipulations, especially regarding controls to the vocalic environment of each stop. Acoustic analyses of other Dhegiha subfamily Siouan languages could be undertaken to establish whether the contrasts in those languages conform to the typical categorical msec values for

127

voiced, voiceless unaspirated, and voiceless aspirated as it appears Omaha stops do. Finally, a revised methodology based on this study could be applied to specific frequently used Omaha words to determine how stops in these words are pronounced today amongst native speakers.

#### Omaha language: revitalization

Although the focus of the current acoustic analysis is on VOT and the identity of phonetic categories in Omaha, it is important to apply this information in the social environment of the language. Thus, in addition to the experimental question, we pose another, larger picture question: how can information gathered in this experiment be used to contribute to the existing Omaha language revitalization program and Omaha language lessons?

Today, the Omaha reservation lies in Thurston County in northeastern Nebraska and incorporates the towns of Macy and Walthill. As of the 2000 census, there were 5194 people living on the reservation. In 1989, there were around 100 fluent speakers on the reservation and neighboring areas (Rudin 1989:2). Today, there are fewer than 50

(Nebraska Indian Community College UmoNhoN Language Survey 2008). The Nebraska Indian Community College in Macy, NE has implemented, among other programs, an Omaha language curriculum, designed and carried out by Alice Saunsoci, Omaha elder and program director, and Ardis Eschenberg, a linguist who lived in Walthill for the past nine years. The lessons, which the author attended for two weeks, are designed to minimize or eliminate dependency on written materials of all kinds, although pictures and

128

objects are frequently used to scaffold the spoken language input which dominates class time. A written version of the lesson material exists exclusively for use after class for additional practice opportunities. This material consists of a non-IPA orthography deemed most accessible to potential learners of Omaha. Saunsoci and Eschenberg based their pedagogical decisions both on L2 acquisition research, and on the traditionally oral nature of Omaha language learning and use.

There is also a program in the public school in Macy, NE that teaches Omaha. The language and culture program was launched in 1970, and has employed various instruction techniques . . . including immersion, memorization of words and phrases, and publication of student authored stories in English and Omaha (Rudin 1989:1). At the time of the Rudin publication, the Omaha language exposure also included students' writing of stories, which were then translated into Omaha by elders and published at the school print shop. Stories were not typically traditional, and the books were in English as well as Omaha, but the activity bolstered a sense of tribal community and identity, as well as affirmative contact between young and elder members of the tribe (Rudin 1989).

The present study can contribute to the existing Omaha curricula by integrating existing pronunciation teaching methods. Speaking and listening are emphasized, but elders may be dissatisfied with students' accidental uttering of inappropriate words due to mispronunciation. Thus, a very important part of the acquisition process is the ability to demonstrate pronunciation deemed adequate by elders and fluent speakers of the language. However, the perception difficulties that native English speaking Omaha

129

people experience when learning Omaha phonemic contrasts can hinder production deemed appropriate by elders.

Although one NICC instructor spoke only Omaha until age nine, other instructors may need guidance with pronunciation patterns. Throughout the process of this study, many inconsistencies were found in the transcription of individual lexical items. This study can serve as a less disputable standard for regularizing interpretations of the stop series' transcription. Non-native-speaking teachers and students can be aided by this pronunciation guide and standard. Also, if a phonemic category in Omaha includes a spectrum of VOT values which have different phonemic interpretations in English, then perception and production may be more difficult than for categories whose VOT values align approximately with English ones (Weber and Cutler 2004, Best 1995). This insight can help instructors teach communicatively while being aware of potential student difficulties. This, in turn, will help native students pronounce their tribal language in the way their elders prefer, stimulating tribal identity and intergenerational ties, and contributing to the invigoration of Omaha language and culture.





Table 4

Charts by number in support or not in support of the hypothesis, and with inconclusive data.

Support of hypothesis by chart number and analysis

	Support					Do not support										Inconclusive				
individual VOT	2	3	11	14	18	1	4	5	6	7	8	9	10							
indiv. speaker aves.	19	20	21	23	24	12	13	15	16	17										
ave by fluency	25	26	27			22														
by flu./word pos.	28	29	30	31	32	33														
by flu./stress	34	35	36	37	39											38				
by flu./w. pos/stress	40	41	43	45	46	47	48									42	44	49	50	51
by flu./vowel	53	54	56	57	59	60		52	55	58										
by flu./w. pos/stress	79	80	81	83	84	86	87	82	85							61	62	63	64	65
																66	67	68		
																69	70	71	72	73
																74	75	76		
																77	78	88	89	90
																91	92	93		
																94				

Table 5

Average VOT values by word position and language, from Lisker and Abramson (1964).

<b>Word</b>							
<b>pos.</b>	<b>Language</b>	<b>/b/</b>	<b>/p/</b>	<b>/p<sup>c</sup>/</b>	<b>/p<sup>h</sup>/</b>	<b>/d/</b>	<b>/t/</b>
Initial	Armenian	-72		5	51	-107	13
Non-in.	Armenian	-47		7	53		* 10
Initial	Thai	-35		8	37	-53	15
Non-in.	Thai	-66	11		50	-38	8
Initial	Korean			7 22	89		11
Non-in.	Korean			5 13*	75		12
<b>Word</b>							
<b>pos.</b>	<b>Language</b>	<b>/t<sup>c</sup>/</b>	<b>/t<sup>h</sup>/</b>	<b>/g/</b>	<b>/k/</b>	<b>/k<sup>c</sup>/</b>	<b>/k<sup>h</sup>/</b>
Initial	Armenian		35	-66	23		83
Non-in.	Armenian		47	-21	27		76
Initial	Thai		63		15		69
Non-in.	Thai		43		16		74
Initial	Korean	30	100		20	48	125
Non-in.	Korean	22*	78		21	44*	93

Table 6 Average VOT values by word position for Omaha speakers, data from  
this study.

	/b/	/p/	/p <sup>h</sup> /	/d/	/t/	/t <sup>h</sup> /	/g/	/k/	/k <sup>h</sup> /
<b>Word pos.</b>									
Initial	-64	10			17		-47	27	40
Medial	-76	21	25	-57	19	49	-54	30	59

## APPENDIX

### Questionnaire—three way stop contrast in Omaha—IPA To be pronounced and recorded by language helpers

#### Bilabials

1. **Bashté**-t<sup>h</sup>e údo<sup>n</sup> (**strawberries** are good)
2. **pahr**<sup>n</sup>-t<sup>h</sup>e údo<sup>n</sup> (the **quill** is good)
3. **zháp**<sup>h</sup>ahi-t<sup>h</sup>e údo<sup>n</sup> (**rosin-weed** is good)
4. nó<sup>n</sup>**pe**-t<sup>h</sup>e údo<sup>n</sup> (she fears) (**fearing** it is good)
5. **hébe**-t<sup>h</sup>e údo<sup>n</sup> (a **piece** is good)
6. **égip**<sup>h</sup>e-t<sup>h</sup>e údo<sup>n</sup> (it<sup>ø</sup> good **I said it**)
7. **biká**-t<sup>h</sup>e údo<sup>n</sup> (it<sup>ø</sup> good that **he wiped** it)
8. **bthípi**-t<sup>h</sup>e údo<sup>n</sup> (that **I know how** is good)
9. húhú-k<sup>h</sup>e-sháge-i-thixáp<sup>h</sup>i-t<sup>h</sup>e údo<sup>n</sup> (using claws to **pierce** the fish is good)
10. **buthít**<sup>h</sup>e-ak<sup>h</sup>á údo<sup>n</sup> (**proper name?** is good)
11. uthí**pu**-t<sup>h</sup>e údo<sup>n</sup> (**bundling up** is good)
12. ti-áthip<sup>h</sup>**u**-tho<sup>n</sup> údo<sup>n</sup> (the **sweat lodge** is good)

#### Dentals

13. múudada-t<sup>h</sup>é píazhi (to **ache** is bad)
14. níta-t<sup>h</sup>e údo<sup>n</sup> (that he is **alive** is good)
15. wathát<sup>h</sup>a-t<sup>h</sup>e údo<sup>n</sup> (that **they were eating** is good)
16. mó<sup>n</sup>**de**-t<sup>h</sup>e údo<sup>n</sup> (the **bow** is good)
17. **tepí**-tho<sup>n</sup> údo<sup>n</sup> (the **liver** is good)
18. bthát<sup>h</sup>e-t<sup>h</sup>e údo<sup>n</sup> (that **I eat** is good)
19. i<sup>n</sup>dádi-ak<sup>h</sup>a údo<sup>n</sup> (**father** [talking about him] is good)
20. tisó<sup>n</sup>the-t<sup>h</sup>e údo<sup>n</sup> (the **tipi** is good)
21. ó<sup>n</sup>gat<sup>h</sup>i-t<sup>h</sup>e údo<sup>n</sup> (it is good that **we are here**)
22. zhi<sup>n</sup>gá- wéduba ák<sup>h</sup>a údo<sup>n</sup> (**the fourth** child is good)
23. tushpátho<sup>n</sup>-ak<sup>h</sup>a údo<sup>n</sup> (**grandchild** [talking about] is good)
24. áwat<sup>h</sup>uskó<sup>n</sup>ska-t<sup>h</sup>e údo<sup>n</sup> (that they **were in a line with** is good)

## Velars

25. Zhúga-k<sup>h</sup>e údo<sup>n</sup> (the **body** is good)
26. Míkasi-ak<sup>h</sup>a píazhi (**coyote** is bad)
27. zhi<sup>n</sup>gá-ák<sup>h</sup>a údo<sup>n</sup> (**the** [sub.] child is good)
28. skíge-t<sup>h</sup>e píazhi (that itø **heavy** is bad)
29. kehámo<sup>n</sup>zhide-ak<sup>h</sup>a údo<sup>n</sup> (the **diamond-backed turtle** is good)
30. k<sup>h</sup>ethó<sup>n</sup>-di údo<sup>n</sup> (**in the past** it was good)
31. gisíthe-t<sup>h</sup>e údo<sup>n</sup> (**to remember** is good)
32. kigthízha-t<sup>h</sup>e údo<sup>n</sup> (to **wash yourself** is good)
33. ík<sup>h</sup>i-tho<sup>n</sup> údo<sup>n</sup> (the **chin** is good)
34. gudéhi-t<sup>h</sup>e údo<sup>n</sup> (the **hackberry tree** is good)
35. kukúmi-k<sup>h</sup>e údo<sup>n</sup> (the **cucumber** is good)
36. ék<sup>h</sup>u-t<sup>h</sup>e údo<sup>n</sup> (that **I invited him** is good)

## REFERENCES

- Abramson, Arthur, and Kalaya Tingsabadh. 1999. Thai final stops: Cross-language perception. *Phonetica*, 56, 111-122.
- Best, Catherine. 1995. A direct realist view of cross-language speech perception. In Strange, W. (Ed.), *Speech perception and linguistic experience: Issues in cross-language research*. 171-204. Timonium, MD: York Press, Inc.
- Boersma, Paul, and David Weenink. 2007. Praat: Doing phonetics by computer (version 4.6.12) [Software]. Available from <http://www.praat.org/> (27 July 2007).
- Bowern, Claire. 2008. *Linguistic fieldwork: A practical guide*. New York: Palgrave Macmillan.
- Broselow, Ellen. 1987. An investigation of transfer in second language phonology. In G. Ioup & S. H. Weinberger (Eds.), *Interlanguage phonology: The acquisition of a second language sound system*. 261-278. Cambridge: Newbury House Publishers.
- Campbell, Lyle, and Mauricio Mixco. 2007. *A glossary of historical linguistics*. Salt Lake City: University of Utah Press.
- Cho, Taehong, and Peter Ladefoged. 1999. Variation and universals in VOT: evidence from 18 languages. *Journal of Phonetics*, 27, 207-229.
- Cho, Taehong, and James McQueen. 2005. Prosodic influences on consonant production in Dutch: Effects of prosodic boundaries, phrasal accent and lexical stress. *Journal of Phonetics*, 33(2), 121-157. doi:10.1016/j.wocn.2005.01.001
- Curtin, Suzanne; Heather Goad; and Joseph Pater. V. 1998. Phonological transfer and levels of representation: the perceptual acquisition of Thai voice and aspiration by English and French speakers. *Second Language Research*, 14(4), 389-405.
- Davidson, Lisa; Jason Shaw; and Tuuli Adams. 2007. The effect of word learning on the perception of consonant sequences. *Journal of the Acoustical Society of America*, 122(6), 3697-3709.
- Eschenberg, Ardis. 2005. *The article system of Umo<sup>n</sup>ho<sup>n</sup> (Omaha)*. (Unpublished Doctoral Dissertation). State University of New York, Buffalo.

- Hale, Kenneth. 1985. A note on Winnebago metrical structure. *International Journal of American Linguistics*, 51(4), 427-429.
- Hayes-Harb, Rachel. 2007. Lexical and statistical evidence in the acquisition of second language phonemes. *Second Language Research*, 23(1), 65-94.
- Hewitt, John Napoleon Brinton. 1895. James Owen Dorsey. *American Anthropologist*, 8(2), 180-183.
- Ladefoged, Peter. 1990. The revised International Phonetic Alphabet. *Language*, 66(3), 550-552.
- Lisker, Leigh, and Arthur Abramson. 1964. A cross-language study of voicing in initial stops: Acoustical measurements. *Word*, 20, 384-422.
- Maddieson, Ian. 1984. *Patterns of sounds*. Cambridge: Cambridge University Press.
- Miller, Joanne. 2001. Mapping from acoustic signal to phonetic category: Internal category structure, context effects and speeded categorisation. *Language and Cognitive Processes*, 16(5/6), 683-690.
- Nebraska Indian Community College. n.d. The Umo ho (Omaha) Center of Excellence at Nebraska Indian Community College. Retrieved from: [http://www.thenicc.edu/Course\\_sites/default.asp](http://www.thenicc.edu/Course_sites/default.asp)
- Rankin, Robert. (transcription editor). 2008. James Owen Dorsey's The *čegiha* Language. (Unpublished re-transcription). University of Kansas, Lawrence.
- Riney, Timothy; Naoyuki Takagi; Kaori Ota; and Yoko Uchida. 2007. The intermediate degree of VOT in Japanese initial voiceless stops. *Journal of Phonetics*, 35(3), 439-443. doi:10.1016/j.wocn.2006.01.002
- Rudin, Catherine. 1989. Omaha language preservation in the Macy, Nebraska public school. (Paper presented at the Meeting of the Mid-America Linguistics Conference). Cedar Falls, IA.
- Wolff, Hans. 1950. Comparative Siouan I. *International Journal of American Linguistics*, 16(2), 61-66.